



GPT-2:n hyödyntäminen fik- tiosisältöjen käsikirjoittami- sessa

Koneoppiminen elokuvakäsikirjoittajan työkaluna

Vertti Luostarinen

OPINNÄYTETYÖ
Kesäkuu 2020

Media-alan tutkinto-ohjelma
Käsikirjoittaminen, emerging media

TIIVISTELMÄ

Tampereen ammattikorkeakoulu
Media-alan tutkinto-ohjelma
Käsikirjoittaminen, emerging media

LUOSTARINEN, VERTTI:

GPT-2:n hyödyntäminen fiktiosisältöjen käsikirjoittamisessa
Koneoppiminen elokuvakäsikirjoittajan työkaluna

Opinnäytetyö 74 sivua
Kesäkuu 2020

Nykyään kuka tahansa voi hyödyntää tekoälyä sisältöjen luomiseen, ja välineistä on tullut entistä tehokkaampia. GPT-2 -algoritmi kykenee tuottamaan tekstiä missä tahansa lajityypissä. Tässä opinnäytetyössä tutkittiin, minkälaisia käsikirjoituksia se voi tuottaa. Samalla pohdittiin, mitä hyötyä GPT-2:lla ja sen kaltaisilla koneoppimismalleilla voisi olla käsikirjoittajalle. Tutkimuksessa on uutta dialogisen suhteen olettaminen kirjoittajan ja tekoälyn välille: tekoälyn itsenäisesti luomia tekstejä ei rinnasteta sellaisenaan ihmisen luomiin. GPT-2:n kanssa tehtiin kaksi koetta, joissa tutkittiin, millaista dialogia ja kokonaisia käsikirjoituksia sillä voi saada aikaan. Koetulosten arvioinnissa keskeisinä tutkimuskysymyksinä olivat, miten käsikirjoittajan on sovellettava työnkulkuaan tekoälyn kanssa työskennellessä ja minkälaista uutta sisältöä tekoäly voi tuottaa.

Kokeiden lopputuloksia arvioitiin toisaalta käsikirjoitusten ulkoisten, toisaalta sisällöllisten piirteiden toteutumisen näkökulmasta. Ulkoisilla piirteillä tarkoitetaan sitä, miten hyvin generoidut tekstit istuvat käsikirjoitusformaattiin. Sisällöllisillä piirteillä tarkoitetaan tekstin sisäistä koherenssia, sekä sen kykyä ilmentää draamallisia jännitteitä merkityksellisillä tavoilla. Kokeiden tuloksena ilmeni, että GPT-2 kykenee odotetusti käsikirjoitusten generointiin tarkkuudella, johon aikaisemmat mallit eivät ole yltäneet. Käsikirjoittajalle tekoälystä on tällä hetkellä hyötyä ennen kaikkea ideointivälineenä, sekä ihmisjärjelle vieraalta tuntuvan sisällön tuottamisessa.

Tutkimustulokset saavat pohtimaan, millaisia odotuksia asetamme tekoälylle jatkossa. Tekoäly ei tule yleistymään media-alalla ennen kuin sitä aletaan lähestyä sen omilla ehdoilla sen sijaan, että sille jatkuvasti asetettaisiin epärealistisia tavoitteita. Myös suhtautumisen tekoälyyn työskentelyvälineenä on muututtava dialogisemmaksi.

Seuraava vaihe olisi tässä opinnäytetyössä esitellyn työnkulun soveltaminen käytäntöön uudenlaisten tekoälyavusteisten fiktiosisältöjen käsikirjoittamisessa. Tekoälyn herättämä vierauden tunne on niin vahva, että se itsessään voisi synnyttää uuden elokuvan lajityypin, algoritmisen kauhun.

ABSTRACT

Tampere University of Applied Sciences
Bachelor of Culture and Arts
Screenwriting, Emerging Media

LUOSTARINEN VERTTI:

The use of GPT-2 in the screenwriting process of fictional content
Machine learning as a screenwriting tool

Bachelor's thesis 74 pages
June 2020

Nowadays anyone can use artificial intelligence for content creation because the necessary tools are more powerful than ever. The GPT-2 algorithm can produce text in practically any style. The purpose of this study was to examine what kind of fictional screenplays the algorithm can produce and what benefits a screenwriter can gain by using machine learning applications such as GPT-2. As a new approach, a dialogical relationship between the writer and AI was assumed: the texts created independently by AI are not compared as they are to the ones a human would write. Two experiments were conducted with GPT-2: one that involved generating only dialogue and another that involved generating entire screenplays. While assessing the results of the experiments, the key questions were: how does a screenwriter need to adjust his or her workflow while using AI and what kind of new content can AI help to usher in?

The results of the experiments were analysed both in terms of structure and content. In terms of structure, it was examined how well the texts adhered to the established screenplay format. In terms of content, it was analysed how well the texts maintain internal coherence and whether they manage to express dramatic tension in meaningful ways. The results showed that, as expected, GPT-2 can generate screenplays with unmatched accuracy when compared to earlier machine learning applications. For a screenwriter, AI can be useful as a tool for generating ideas at the early stages of screenwriting and later in the process as a means to invoke the feeling of uncanniness in the reader/viewer.

The findings indicate that it is vital to consider which expectations will be set on AI in the future. AI will not become more common in the field of media unless it is approached on its own terms. It is not conducive to set unattainable goals for its use. The approach to AI as a content creation tool must also become more dialogical.

The next step would be to apply the workflow presented in this thesis to practice and use it to create new kinds of AI-assisted fictional content. The alienating feeling that AI invokes in itself is so strong that it could generate a new genre of cinema: algorithmic horror.

Key words: film, cinema, movies, screenwriting, artificial intelligence, GPT-2

SISÄLLYS

1	JOHDANTO	5
2	KONEOPPIMINEN JA NLP	11
2.1	Koneoppiminen	11
2.1.1	Keinotekoiset neuroverkot	11
2.1.2	Syväoppiminen	13
2.1.3	Transfer Learning, eli siirto-oppiminen	14
2.2	NLP ja kieliteknologia	15
2.2.1	NLP:n historiaa lyhyesti	15
2.2.2	Syväoppiminen NLP:ssä ja GPT-2	16
2.3	Olemassa olevia tekoälysovellutuksia media-alalla.....	18
2.3.1.	Musiikki.....	18
2.3.2	Interaktiivinen media.....	19
2.3.3	Kuvien generointi.....	20
2.3.4	Tarinagenerointi ja kirjoittaminen.....	21
3	TUTKIMUS: KOKEET JA NIIDEN TULOSTEN ARVIOINTIA	24
3.1	Dialogin generoiminen	27
3.1.1	Toteutuskeinot ja generointi	28
3.1.2	Lopputulosten arviointia.....	29
3.1.3	Sovellutusmahdollisuudet.....	42
3.2	Kokonaisen käsikirjoituksen generoiminen	45
3.2.1	Toteutuskeinot ja generointi	45
3.2.2	Lopputulosten arviointia.....	46
3.2.3	Sovellutusmahdollisuudet.....	60
4	POHDINTA	64
	LÄHTEET.....	69

1. JOHDANTO

Algoritmilla tarkoitetaan mitä tahansa toimintaohjetta. Tietokone seuraa aina jotakin ohjetta; toisin sanottuna tarvitaan aina algoritmi, jotta tietokone tekee jotain. Ihmisen toimintaa imitoivia algoritmeja kutsutaan tietojenkäsittelyssä tekoälyksi.

Laveammin tekoälyllä tarkoitetaan sellaista ihmisen luomaa mekaanista prosessia, jonka me ihmiset miellämme älykkääksi. Ongelmana tässä on, että se mitä pidämme älykkäänä, on muuttunut vuosien varrella. Tekoälyn ytimenä 50 vuotta sitten pidetyt käsin koodatut laskennalliset prosessit ovat osa nykyisen tietojenkäsittelyn perustaa. Termi tekoäly on siirtynyt merkitsemään kompleksisempia, vaikeammin määriteltäviä prosesseja, esimerkiksi koneoppimista. (Helsingin yliopisto & Reaktor. n.d.)

Määriteltiin tekoäly miten tahansa, oleellista on se, että suhteutamme sen aina siihen, mitä kuvittelemme ihmisen pystyvän tekemään. Ihminen ja hänen kognitiivisten toimintojensa jäljitteleminen ovat olleet universaali mitta tekoälyn edistykselle. Tieteiskirjailijoita on iät ja ajat kiehtonut se ontologinen dilemma, jonka tämä rinnastus tuottaa: onko joku päivä mahdollista tuottaa niin tarkka kopio ihmisestä, että tekoälyn ja inhimillisen älyn välinen raja muuttuu merkityksettömäksi? Historioitsija Yuval Noah Harari tiivistää kirjassaan *Homo Deus – Huomisen lyhyt historia* (2015, 329) juuri alkaneen vuosituhantemme materialistisen ideologian seuraavasti:

1. Organismit ovat algoritmeja. Jokainen eläin – mukaan lukien *Homo sapiens* – on kokoelma orgaanisia algoritmeja, jotka luonnonvalinta on muovannut miljoonien evoluutiovuosien aikana.
2. Algoritmisten laskutoimitusten tuloksiin eivät vaikuta materiaalit, joista laskin on rakennettu. Rakennettiinpa helmitaulu puusta, raudasta tai muovista, kaksi helmeä ynnä kaksi helmeä on neljä helmeä.
3. Näin ei ole mitään syytä olettaa, että orgaaniset algoritmit kykenisivät tekemään asioita, joita epäorgaaniset algoritmit eivät koskaan pystyisi jäljittelemään tai tekemään paremmin. Niin kauan kuin laskelmat pitävät paikkansa, onko sillä väliä, onko algoritmit tehty hiilestä vai piistä?

Tekoälyn teoreettinen potentiaali on mittaamaton, mutta käytännössä olemme vielä kaukana tieteiskirjailijoiden visioista. Ei tarvitse kuin yrittää kirjoittaa pelkällä

automaattisella tekstinsyötöllä huomatakseen, että ihminen on yhä optimaalinen vaihtoehto useimmissa luovissa tehtävissä. Viime aikoina suuryritykset ovat alkaneet spekuloida, että tekoälykehityksessä ollaan törmäämässä seinään. Raja, mihin tekoäly pystyy ja mihin ei, on piirtynyt selkeämmin (Knight 2019). Tekoälyn suitsutus on alkanut jopa herättää turhautumista. Startup-yrittäjä Sami Kuusela kuvailee tekoälystä teoretisoivia ihmisiä ”käsienheiluttelijoiksi” (Kuusela & Kaanpää 2019):

Käsienheiluttelija on se, joka menee koodaajan viereen vaahtomaan ja vaatimaan asioita, joita on mahdotonta toteuttaa. Se pystyy heiluttelemaan käsiä, koska ei tee itse mitään. Käsienheiluttelija nousee usein myös lavalle vetämään keynoteja, joissa se kertoo tekoälyn mahdollisuuksista tai siitä, miten kaikki tulee muuttumaan huomenna. Mutta se ei oikeasti tiedä.

Oma kehitykseni tekoälyn kanssa on edennyt pikemminkin käytännön tasolta kohti ”käsienheiluttelua”. Aloin kirjoittaa tekoälyavusteista fiktiota jo lukiossa vuonna 2014. Tuohon aikaan välineitä sitä varten oli saatavilla vähemmän ja tekninen osaamiseni oli puutteellista. Turvauduin Nokia-puhelimeni automaattiseen tekstinsyöttöön luodessani tarinoita. Lukion jälkeen kirjoitin Vaeltaja-nimisen tie-teissarjan, jossa hyödynsin kehittämäni työnkulkua ensimmäistä kertaa työkontekstissa. Opintojeni sivussa olen pyrkinyt syventämään teoreettista osaamistani ja laajentamaan tietotaitoani myös visuaalisen sisällön generointiin: esimerkiksi syksyllä 2019 tein koulutyönä Valtteri Bottas -deepfaken.

Viime vuosina on tapahtunut valtava tekninen murros, kun tekoälysovelluksista on vähitellen tullut valtavirtaa: tätä nykyä käytännössä kuka tahansa voi operoida koneoppimisalgoritmeja kotikoneellaan tai pilvipalvelun välityksellä. Tekoäly on tämän myötä raivaamassa tietään kaikille elämänaloille. Sovellutukset ovat helpokäyttöisempiä, parempia ja helpommin saatavilla. Henkilökohtainen tavoitteeni on ollut löytää tapoja, joilla tekoälyä voisi hyödyntää media-alalla. Osaajia, jotka taitaisivat sekä algoritmit että sisällöntuotannon, on todella vähän.

Tekstingeneroinnin alaa on hiljattain uudistanut OpenAI:n GPT-2 -algoritmi, joka kykenee itsenäisesti tuottamaan tekstiä miltei missä tahansa lajityypissä. Uutta GPT-2:ssa on sen yleisluontoisuus: sitä kykenee sellaisenaan soveltamaan miltei tehtävään kuin tehtävään: niin ruokaohjeisiin kuin uutisiin (Radford ym. 2019).

Minua alkoi kiinnostaa, miten hyvin GPT-2 voisi soveltua fiktiosisältöjen käsikirjoittamiseen. Olen edellisen kerran pohtinut tekoälyavusteisen käsikirjoittamisen ongelmia ja työnkulkua silloin, kun käytössäni on ollut vain oma puhelimeni. Nyt tilanne on aivan erilainen kuin kuusi vuotta sitten.

Tämän tutkimuksen tavoitteena on selvittää, missä määrin tekoälyn generoimaa sisältöä voidaan integroida käsikirjoitusprosessiin. Mitkä ovat tällaisen integroinnin mahdolliset edut ja haitat? Miten käsikirjoittajan täytyy muuttaa työskentelytapojaan tekoälyn kanssa ja millaisia uusia sisällöllisiä mahdollisuuksia tekoälyn hyödyntäminen avaa? Pyrin löytämään vastauksen näihin kysymyksiin kokeilemalla, miten GPT-2:ta voisi soveltaa fiktiokäsikirjoituksen kirjoitusprosessissa. Pohdin sitä, miten hyvin GPT-2:n teksti vastaa käsikirjoitusta niin ulkoisilta kuin sisällöllisiltä piirteiltään. Tutkin tätä sekä pelkän dialogin että kokonaisten käsikirjoitusten näkökulmasta. Kuvaan myös, millaisia teknisiä muuttujia generointiprosessiin liittyy ja miten ne voivat vaikuttaa lopputulokseen.

Käytän GPT-2:ta Googlen Colaboratory -pilvipalvelussa, jossa kuka tahansa voi pyörittää koodia Googlen palvelimilla. Koulutan algoritmia eri kokoisilla korpukilla, eli tekstidataseiteillä. Ne koostuvat pääasiasiassa yhdysvaltalaisen toimintaelokuvien käsikirjoituksista, jotta lopputulos olisi mahdollisimman yhtenäinen. Suoritan kaksi koetta, joista ensimmäisessä tutkin GPT-2:n kykyä generoida dialogia, kun se reagoi ihmisen syötteeseen. Toisessa kokeessa tutkin sen kykyä generoida kokonaisia käsikirjoituksia itsenäisesti.

Tämän työn tutkimusosa koostuu valikoimieni generointitulosten arvioimisesta. Tekstinäytteitä tutkiessa otteeni on laadullinen. Laadullisessa tutkimuksessa keskitytään tutkittavan asian kommunikoimiin merkityksiin ja vivahteisiin, usein yleistettävyyden kustannuksella (Tuomivaara 2005). Olen toisin sanottuna valinnut sellaisia näytteitä, joista ilmenee jotakin uutta tutkimuskysymysten valossa. Ne eivät ole edustava otanta, vaan valikointiprosessini tulos.

Uusi teknologia luo aina uudenlaisen työnkulun. Pohdinkin paitsi sitä, millaista sisältöä koneoppimismallit pystyvät tällä hetkellä tuottamaan, myös sitä, millaisen sisällön tuottamiseen ne parhaiten soveltuvat ja miten kirjoittajan on mukautet-

tava omia prosessejaan työskennellessään tekoälyn kanssa. Nähdäkseni kirjoittajan kyky mukautua uuteen teknologiaan on aivan yhtä tärkeää kuin teknologian kyky vastata kirjoittajan tarpeisiin. Miten kirjoittajan on muutettava suhtautumistaan kirjoitusprosessiin, kun mukana on algoritmi? Orientoitumistapaa muuttamalla voitaisiin kenties jatkossa välttyä niiltä ongelmilta, joita on ollut aikaisemmissa yrityksissä tuoda tekoäly osaksi media-alan sisällöntuotantoa.

Olen löytänyt luovien alojen parista paljon Kuuselan ”käsienheiluttelijoita”, mutta vähän ihmisiä, jotka ovat aidosti soveltaneet tekoälyä. Suomalainen markkinointiyritys Kurio piti vastikään Cannesin elokuvajuhlilla työpajan tekoälyn hyödyntämisestä luovissa sisällöissä, eikä se ollut ainoa, joka puhui tekoälystä kyseisellä festivaalilla (Alkula 2019). Jää nähtäväksi, etenevätkö tekoälysovellutukset ikinä konsulttifirmoista ja keynoteista tekijöille asti. Asennoituminen tekoälyä kohtaan on muuttumassa, mutta eri asia on, uskalletaanko sitä aidosti soveltaa.

Toistaiseksi tuotantoon asti edenneitä esimerkkejä on niukasti. Eniten huomiota viime vuosina on herättänyt lyhytelokuva Sunspring, jonka tieteiselokuvien pohjalta koneoppimisalgoritmeilla generoitu käsikirjoitus on näennäisen järjetön (Marks 2018). Sen saama vastaanotto oli parhaimmillaankin laimean huvittunut. Sunspring jäänee anekdootiksi niin elokuvan kuin tekoälynkin historiassa, sillä sen lähtökohta oli liian kunnianhimoinen: tekoälyn ei voinut vuonna 2018 olettaa osaavan luoda mitään mielekästä muutaman käsikirjoituksen pohjalta.

Se, että generoinnin tuloksista ei ole olemassa onnistuneita esimerkkejä, voi osaltaan selittää sitä, miksi käsikirjoittajat eivät ole ottaneet tekoälytyökaluja vastaan erityisen lämpimästi. Kun käsikirjoitussivusto Blacklist alkoi tarjota automatisoitua käsikirjoitusten kommentointipalvelua ScriptBookia, sivuston käyttäjien viha oli niin valtava, että Blacklist veti palvelun pois sivuiltaan heti seuraavana päivänä (Leonard 2017).

Sunspring -esimerkistä voi havaita, että tekoälyn soveltaminen sellaisenaan narratiivien rakentamisessa tavalla, joka tyydyttäisi valtavirtaelokuvan vaatimuksia, on käytännössä mahdotonta. Blacklist -esimerkki taas osoittaa, että elokuva-ala suhtautuu nihkeästi tekoälyavusteisuuteen ja torjuu automaation, jopa sellaisen, jolla on vähän tekemistä itse luovan prosessin kanssa. Suhtautuminen alalla on

tavallaan paradoksaalista: tekoälyn käytännön tuloksia pilkataan, mutta silti sen potentiaali pelottaa. Ei lieneäkään sattumaa, että käsikirjoittajien tekoälyvisiot ovat usein dystooppisia.

Kuten edellä totesin, oma lähtökohtani on, että sen sijaan, että tekoälyn saavutuksia punnittaisiin sen perusteella, miten ne suhteutuvat ihmisen luomuksiin, voitaisiin arvioida, mitä täysin uutta annettavaa niillä voisi olla. Tekoäly ei välttämättä vastaa nykyisiin tarpeisiimme, mutta kenties kehityksen myötä tarpeet itsessään saattavat muuttua. Jokainen teknologinen innovaatio on ajanut eteenpäin myös tarinankerrontaa. Jos vertaa tämän päivän striimauspalveluiden tv-kerrontaa siihen, millaista se oli broadcast-ohjelmissa vielä 20 vuotta sitten, ero on kuin yöllä ja päivällä. Teknologian kehittyessä normimme ja odotuksemme muuttuvat. Muutamana kymmenen vuoden päästä saattaa olla jo vaikeaa kuvitella, miten ikinä pärjäsimme ilman tekoälyavusteista sisällöntuotantoa. Tutkimukseni loppupuolessa pohdinkin, miltä tekoälyavusteinen tarinankerronta voisi näyttää, ja millaisia uusia mahdollisuuksia se luo sisällöntuottajalle. Kun tekoälyä lähestytään sen omilla ehdoilla, monet sille luontaiset ominaisuudet, joita aikaisemmissa tutkimuksissa ja julkisessa diskurssissa on pidetty heikkouksina, muuttuvatkin sen vahvuuksiksi.

Facebookin Yann LeCun on pitkään puhunut siitä, että tekoälyn tuotokset jäävät vajaiksi, koska siltä puuttuu ns. ”world model”, malli sitä ympäröivästä maailmasta (Lu 2018). Tekoäly ei siis kykene asettamaan löydöksiä osaksi suurempaa kokonaisuutta ja maailmankuvaa. Tämä voi olla paitsi haitta, myös etu, sillä se mahdollistaa sen, että tekoäly löytää sellaisia yhteyksiä, joita ihmisen skeemoihin ja konkreettiseen maailmaan kiinnittynyt ajattelu ei ikinä huomaisi. Luovuus ei ole uusien ideoiden synnyttämistä tyhjästä, vaan odottamattomien yhteyksien löytämistä.

Luovuutta on vaikea arvioida olematta jossakin määrin subjektiivinen. En aio altistaa generoituja tekstejä esimerkiksi klassisille aristotelisille tarina-analyysille. Olemassa olevat keinomme arvottaa tarinoita ovat joka tapauksessa tarkoitettuja ihmisen luomien tekstien arviointiin, eivätkä suoranaan soveltuisi koetulosteni vertailuun.

Aiempaa tietämystä tekoälystä ei vaadita tämän opinnäytetyön ymmärtämiseen. Lukion pitkän matematiikan oppimäärän ja ohjelmoinnin alkeiden ymmärtämisestä on taatusti hyötyä, mutta esitän eri tekoälymallit niin yksinkertaisesti kuin suinkin mahdollista. Tekoälyn toimintaa on helppo myös yliyksinkertaistaa, sillä siitä on mutkatonta käyttää sellaisia termejä kuin vaikkapa "ajattelu" tai "muisti", vaikka tosiasiaassa kyse on matemaattisista prosesseista, jotka voisi ilmaista paljon tarkemmin kaavoilla. Kannattaakin siis työtäni lukiessa pitää mielessä, että vaikka koneoppiminen muistuttaa osittain ihmisen oppimista, sen käyttämät prosessit ja mekanismit ovat melko toisenlaiset.

Tämä opinnäytetyö on kirjoitettu sisällöntuottajan näkökulmasta, ja tämä heijastuu myös käyttämissäni välineissä, jotta löydöksistäni voisi olla potentiaalisesti apua kenelle tahansa alalla työskentelevälle. Työn on tarkoitus toimia esimerkiksi käsikirjoittajalle, joka tahtoo hyödyntää tekoälyä kirjoittamisprosessissaan. Yleistyäkseen tekoälysovellutusten on pakko olla helppokäyttöisiä: jos olettaisimme kirjoittajien osaavan koodata, se olisi sama kuin olettaisimme kuvaajien osaavan valmistaa kamerasensoreita tai leikkaajien kehittävän kuvanpakkausalgoritmeja. Googlen Colaboratory ei vaadi tehokasta tietokonetta pyöriäkseen ja käyttämiäni malleja voi seurata melko vähäisellä koodaamiskokemuksella. Tämä tarkoittaa myös sitä, ettei tutkimukseni edusta, eikä voikaan edustaa teknisen edistymisen terävintä kärkeä. Vaikeammilla prosesseilla saisi aikaan paljon parempia lopputuloksia.

2. KONEOPPIMINEN JA NLP

2.1 Koneoppiminen

Koneoppimisella tarkoitetaan sellaisia tekoälyn sovellutuksia, joissa tekoälylle annetaan dataa ja se laitetaan oppimaan itsenäisesti sen pohjalta (Marr 2016). Toisin sanottuna koneoppimisessa ohjelmoija ei ole kirjoittanut lukkoon lyötyjä algoritmeja, vaan tekoälyä on ohjeistettu oppimaan esimerkeistä (Roos 2019).

Nykyaikana datan kerääminen ja säilöminen on helppoa ja yksinkertaista, mutta sen prosessointi vaatii valtavasti työtä. Koneoppiminen on hyödyllistä etenkin silloin, kun datan pohjalta muodostettavien vaihtoehtojen määrä on liian laaja otettavaksi huomioon perinteisessä ohjelmoinnissa. Koneoppimiselle luonteenomaista on myös, että tietokone parantaa suorituskyykyään sitä mukaa, kun se saa uusia ja tarkempia esimerkkejä (Roos 2019). Datan määrä ja tarkkuus ovat tulosten kannalta tärkeitä tekijöitä, samoin käytetty koneoppimismalli.

Koneoppimista hyödynnetään tällä hetkellä ennen kaikkea data-analyysien ja yhteenvetojen tekemiseen. Yrityksissä sitä hyödynnetään eniten myynnissä ja markkinoinnissa, esimerkiksi markkinoinnin kohdentamisessa (Columbus 2019). Nykyisin miltei jokainen edistyksellinen tekoälysovellutus hyödyntää tavalla tai toisella koneoppimista, erityisesti niin kutsuttua syväoppimista (Hardesty 2017).

2.1.1 Keinotekoiset neuroverkot

Keinotekoiset neuroverkot ovat koneoppimisessa hyödynnettäviä algoritmeja, joissa on pyritty aivofysiologisen tutkimuksen pohjalta jäljittelemään sitä, miten ihmisen tiedonkäsittelyprosessit rakentuvat (Honkela 1996). Kuten ihmisen aivot, myös keinotekoiset neuroverkot prosessoivat useita asioita rinnakkain. Neuroverkot koostuvat tuhansista tai miljoonista prosessointiyksiköistä, jotka ovat kyöksissä toisiinsa (Hardesty 2017). Toisin kuin perinteiset algoritmit, neuroverkot

noudattavat ns. sumeaa logiikkaa. Oppiminen tapahtuu keinotekoisten hermosolujen välillä olevien liitosten, keinotekoisten synapsien, voimakkuuksien vaihteluina, eikä yksinkertaisina binäärisinä asiantilojen muutoksina.

Neuroverkkoja on siis sekä orgaanisia että keinotekoisia, mutta tämän tekstin kontekstissa, kuten alan tutkimuksessa yleensäkin on tapana, en ole katsonut aiheelliseksi mainita verkkojen keinotekoisuutta joka kerta erikseen.

Jokaisella keinotekoisella synapsilla on ns. synaptinen paino, joka määrittää kertoimen, jolla se arvottaa lävitseen kulkeutuvaa dataa. Keinotekoiset neuronit summaavat eri puolilta saamiaan signaaleja ennen kuin lähettävät niitä vuorostaan eteenpäin synapseja pitkin. Jos summattu luku ei ylitä ennalta asetettua aktivaatiokynnystä, dataa ei siirry eteenpäin. Kun neuroverkkoja koulutetaan, verkko opettelee näitä painoja ja kynnyksiä, kunnes se tuottaa halutunlaisia tuloksia. (Hardesty 2017).

Normaaleissa neuroverkoissa data kulkee vain yhteen suuntaan. Jos kuitenkin täytyy tuottaa dataa, johon vaikuttavat edelliset tulokset, tällainen laskenta ei enää riitä. Esimerkiksi automaattisessa tekstinsyötössä on tärkeää, että neuroverkko seuraavaa sanaa ennustaessaan katsoo useamman sanan verran taaksepäin; muuten se todennäköisesti jää vain junnaamaan paikalleen (Gudigandula 2019). RNN:t (Recurrent neural networks), eli takaisinkytkettyvät neuroverkot prosessoivat nimensä mukaisesti uudestaan sisään tulevana datana sitä, mitä ne ovat hetkeä aiemmin itse tuottaneet (Nicholson n.d.). Tavallaan voisi sanoa, että niillä on muisti, jonka avulla ne voivat hyödyntää aiemmin oppimaansa. Tärkeää ei kuitenkaan aina ole vain tarpeellisen muistaminen vaan tarpeettoman unohtaminen, ja onkin olemassa takaisinkytkettyviä verkkoja, jotka on suunniteltu myös tätä silmällä pitäen.

Harvoin on tiedossa, mitä mikäkin synapsi tarkalleen tekee tai minkälaisista attribuuteista sen analyysi koostuu. Neuroverkko voi esimerkiksi osata tunnistaa erinomaisesti kasvoja, mutta on vaikeampaa tietää, minkä osatekijöiden perusteella se erottelee kasvot toisistaan. Ihminen osaa tietoisesti käsitteellistää ja purkaa kasvoista tunnistettaviksi ominaisuuksiksi vaikkapa silmien värin tai hiustyylin,

mutta neuroverkossa kaikki tekijät on esitetty numeerisesti tavalla, josta on mahdollonta löytää vaikkapa yhtä tiettyä ”silmienväri-synapsia”. Neuroverkon toiminta on redusoitavissa numeroiksi, mutta harvoin käsitteiksi. Siksi neuroverkkoja on pitkään nimitetty ns. ”mustiksi laatikoiksi”. Tietoa menee sisään ja tulee ulos, mutta varsinaista prosessia on vaikea hahmottaa (Alain & Bengio 2018). Prosessia on yritetty kääntää, jotta nähtäisiin konkreettisesti, minkälaisista osatekijöistä koneen hahmottaminen koostuu. Tästä hyvä esimerkki on Googlen Deep Dream -projekti, jossa neuroverkko on laitettu etsimään ja täydentämään kasvoja kuviin, joissa niitä ei esiinny (Mordintsev, Olah & Tyka, 2015). Deep Dream -kuvat ovat painajaismaisia, mutta niistä voi havaita, mitä Googlen algoritmi yrittää etsiä kasvoja tunnistessaan.

Koska neuroverkkojen prosessit ovat monesti käytännössä läpinäkymättömiä, oikeiden parametrien löytäminen omiin uniikkeihin käyttötarkoituksiin on monesti enemmän taidetta kuin tiedettä. Parhaat ratkaisut löytyvät kokeilulla ja kokemuksella. Neuroverkot ovat arvaamattomia ja näennäisen pienillä muutoksilla voi olla valtavia, ennustamattomia seurauksia. Myös koneoppimiselle ominainen runsas hajonta saattaa vaikeuttaa optimaalisten parametrien löytämistä. Aivan pelkkää arvailuakaan niiden parissa työskentely ei kuitenkaan ole. Internetistä löytyy olemassa olevia esimerkkejä ja tutoriaaleja, joita seuratakseen ei tarvitse olla asiantuntija.

2.1.2 Syväoppiminen

Kun keinotekoisista neuroverkoista alettiin ensi kerran puhua 1940 -luvulla, kognitiivinen ja neuropsykologia ottivat ensiaskeleitaan. Käsitteen lanseerasivat Warren McCullough ja Walter Pitts Chicagon yliopistossa. He yrittivät osoittaa, että ihmisen tiedonkäsittely on teoriassa mallinnettavissa tietokoneella (McCullough & Pitts 1943). Käsite esiintyi tietojenkäsittelytieteessä, psykologiassa ja neurotieteessä vuosikymmenien varrella, mutta teknologia ei pitkään aikaan sallinut kovin kompleksisten verkkojen mallintamista (Anderson & Rosenfeld 1998).

2010-luvulla käsite nousi uudelleen tieteellisen keskustelun keskiöön, kun tietokoneiden näytönohjainten kehitys oli edennyt pisteeseen, jossa ne kykenivät prosessoimaan kymmeniä ja taas kymmeniä neuroverkkokerroksia samanaikaisesti. Kerrosten määrän lisääntyessä oppiminen syventyi ja alettiin puhua ”syväoppimisesta”. Tänä päivänä miltei jokaisen tietokoneen näytönohjain kykenee syväoppimiseen tehokkuudella, joka ei vielä vuosituhannen vaihteessa ollut kuviteltavissakaan (Hardesty 2017).

Kerrosten määrän lisääntyessä ovat lisääntyneet paitsi neuroverkoille annettujen tehtävien kompleksisuus, myös tulosten tarkkuus (Telgarsky 2016). Monet kehittyneet sovellutukset hyödyntävät useita, erikoistuneita neuroverkkotalgoritmeja, joiden syvyyttä ja järjestystä on mahdollista myös varioida tarpeen ja laskentatehon mukaan. Syvyys ei automaattisesti tarkoita aina parempia tuloksia. Kaikki riippuu käytössä olevien välineiden tarkoituksenmukaisesta soveltamisesta.

2.1.3 Transfer Learning, eli siirto-oppiminen

Ihmiset ovat normaalisti hyviä siirtämään samat taidot kontekstista toiseen. Kun on oppinut pelaamaan yhtä videopeliä, seuraavan opettelu on helpompaa. Ihmisen, joka jo nuorella iällä on oppinut puhumaan vierasta kieltä, on helpompi oppia puhumaan muitakin kieliä. Perinteisessä koneoppimisessa kerrytetty osaaminen ei samalla tavalla siirry tehtävästä toiseen. Koneoppimisalgoritmit ovat hyvin erikoistuneita tietyn tyyppisiin tehtäviin, eikä samaa koodia voi sellaisenaan hyödyntää toisessa kontekstissa.

Transfer learningissa, eli siirto-oppimisessä pyritään siihen, että algoritmi kykenee itsenäisesti hyödyntämään yhdessä tehtävässä oppimaansa toisessa tehtävässä (Sarkar 2018). Siitä povataan seuraavaa läpimurtoa alalla (Ng 2016). Yksi onnistuneimpia siirto-oppimiseen pohjautuvia kokeiluja on OpenAI:n kehittämä GPT-2, jota käytetään tässä tutkimuksessa. Siirto-oppimisen mahdollistama muokautuvuus on tärkeä tekijä sille, että malli on ollut niin menestyksenkäs. GPT-2 hyödyntää siirto-oppimisen alle lukeutuvaa multitask learningia, eli monitehtäväoppimista, jossa malli opettelee yhtä aikaa useita eri tehtäviä. Tämä poikkeaa perinteisestä siirto-oppimisesta, jossa mallin oppimista ei alkuun ohjata millään tietyllä tehtävällä (Sarkar 2018).

2.2 NLP ja kieliteknologia

NLP, eli Natural Language Processing tutkii ihmisten luonnollisen kielen ja tietokoneiden vuorovaikutusta. Tekoälyn saralla tärkeitä tutkimuskohteita ovat esimerkiksi luonnollisen kielen generointi (NLG) ja puheen- ja tekstintunnistus. Suomessa NLP kuului alun perin tietokonelingvistiikan oppiaineeseen, joka kääntyi 2000-luvulla muotoon kieliteknologia. Helsingin yliopiston mukaan kieliteknologia tieteenalana tutkii sitä, miten kieli jäsentyy tietokoneelle ymmärrettävään muotoon (Language Technology n.d.). Termit tarkoittavat siis käytännössä melkein samaa asiaa, vaikka kieliteknologia ei olekaan aivan suora käänös NLP:stä.

2.2.1 NLP:n historiaa lyhyesti

1600-luvulla valistusfilosofi Gottfried Leibniz kehitti ajatuksen ”mielen laskukoneesta”, joka kykenisi muuntamaan ajatuksen kuin ajatuksen symboliseen muotoon, ”universaalille kielelle”, ja laskemaan täydellisen, rationaalisen vastauksen sille annettuun kysymykseen (Roinila 1996). Ajatus oli järjen aikakaudellakin radikaalin optimistinen; jopa äärimmäisestä rationalismistaan tunnettu Descartes oli irtisanoutunut ajatuksesta ”puhuvasta ihmiskoneesta”, joka vastaisi kuin ihminen joka kerta riippumatta kysymyksestä.

Alan Turing esitti 1950 tekstissään Computing Machinery and Intelligence kuuluisan ajatuksen Turingin testistä, jossa juuri kieli ja kommunikaatiokyky näyttäytyivät tekoälyn uskottavuuden mittana (Turing 1950). Testissä ihmisvastaanottaja kommunikoi oven ali syötettävien lappujen kautta koneen kanssa näkemättä tätä. Testi onnistuu, kun ihminen ei enää erota, että toisella puolella ovea on kone. Turingille oli tärkeää inhimillisen kielen ja toiminnan representoiminen uskottavasti, eikä se, onko koneella tietoisuutta. Turing tunnetaan toisessa maailmansodassa saksalaisten koodien murtamisesta Colossus -tietokoneella, joka kykeni purkamaan numeroita sanoiksi hyödyntäen kielen numeerista analyysia (Copeland n.d.). Laite oli myös ensimmäisiä, joka purki kieltä koodiksi, Colossuksen tapauksessa 5-bittiseksi binäärikoodiksi.

Kielitiede ja kieliteknologia ovat aina kulkeneet rinnakkain. Tärkein kysymys NLP:ssä on aina ollut, miten luonnollisen kielen syntaksia ohjelmoidaan, eli siis miten ihmisten kielellinen kommunikaatio kaikkine sääntöinen ja nyansseineen saadaan sellaiseen muotoon, jossa tietokone ymmärtää sitä. Kielitieteilijä Noam Chomsky kehitti yksinkertaisen ja tehokkaan hierarkkisen järjestelmän kielen kvantifioinnille ja mallintamiselle jo 1950-luvulla, ns. Chomskyn hierarkian (Chomsky 1956). NLP-tutkimus oli pitkään lähinnä tällaisten mallien siirtämistä formaaliin, matemaattiseen muotoon.

Tammikuussa 1954 Georgetown-IBM-kokeessa onnistuttiin kääntämään täysin koneellisesti useita venäjänkielisiä lauseita englanniksi (Hutchins 2004). Kokeen suorittaneet tutkijat Georgetownin yliopistosta julistivat ratkaisevansa kaikki konekääntämisen ongelmat korkeintaan viidessä vuodessa (Hutchins 2005). Kuten monta kertaa tekoälyn historiassa, odotukset asetettiin liian korkealle. Turingin tai Leibnizin visioiden esteenä oli vielä monia kymmeniä vuosia teknologinen kehitys; kieliopin säännöt ja sanastot piti koodata manuaalisesti järjestelmiin. Se oli aikaa vievää ja työlästä. Lisäksi kehitystyötä rajoitti kielen muovautuvuus ja taipumus paeta tarkkoja määrittelyitä. On vaikeaa jäljitellä asiaa, jonka toimintamekanismia luonnossa ei vielä tunneta kunnolla.

Jo 1960-luvun alussa ymmärrettiin, että kieliteknologia vaatisi sivistyneempää tekoälyä kuin mihin silloinen prosessointiteho riitti. Tutkijat eivät olleet osanneet aavistaa, miten vaikeaa NLP:tä oli soveltaa jopa kaikkein kapeimpiin reaali maailman sovellutuksiin (Jones 2001). Alkeellisia kielenennustamisjärjestelmiä, kuten chatbotteja, eli keskustelujärjestelmiä, ja hakukoneita oli jo 1970-luvulla, mutta kesti pitkään, ennen kuin niistä tuli valtavirtaa.

2.2.2 Syväoppiminen NLP:ssä ja GPT-2

Syväoppimisen läpimurrot nostivat kieliteknologian jälleen esille 2000-luvun taitteessa. Mallit, jotka oppivat suoraan tekstiä analysoimalla alkoivat vähitellen syrjäyttää vanhat järjestelmät. Hakukoneista, puheentunnistuksesta ja konekääntämisestä on sittemmin tullut arkipäivää. Tosin tarve saneistukselle (tokenisation), eli kieliopin opettamiselle tekoälylle, ei ole kadonnut minnekään. Tekoäly on miltei

aina yhä opetettava pilkkomaan sanat oikeista kohdista, jotta se osaisi muodostaa semanttisesti järkeviä kokonaisuuksia. Esimerkiksi suomen kielen sijamuodot ovat yhä Android -puhelinten edistyneelle automaattiselle tekstinsyötölle liian suuri haaste.

2020-luvulla kieliteknologian suurimpia haasteita on yhä Leibnizin ”universaalin kielen” löytäminen. Miten kääntää kaksi eri kieltä rinnasteiseen numeeriseen muotoon? Englanti taipuu koneilta jo Turingin testin läpäisyä hipovilla tavoilla, mutta esimerkiksi arabia jo paljon nihkeämmin, sillä se vaatii aivan erilaista saneistusta ja syntaksia (Farghaly & Shaalan 2009). Myöskään tarve kuratoidulle kielidatalle ja tekstin massa-analyysille ei ole poistunut. Osasyynä englannin kielessä tehdyille läpimurroille on ollut saatavilla olevien aineistojen määrä verrattuna pieniin tai edes keskikokoisiin kieliin. Tämän hetken edistyneimmät mallit tarvitsevat miljoonia tai jopa miljardeja sivuja toimiakseen.

Tämän hetken lupaavin ja kohutuvin hanke on yhdysvaltalaisen OpenAI:n GPT-2, joka julkaistiin erissä vuoden 2019 aikana. Hanke oli kiistanalainen, koska tutkijat tiedeyhteisön sääntöjen vastaisesti julkaisivat GPT-2:n alkuun vain rajattuna mallina, jolle oli opetettu murto-osa koko saatavilla olleesta datasta (Vincent 2019). Tutkijat perustelivat tätä sillä, että he eivät tahtoneet, että mallia väärinkäytetään. GPT-2 ei erota totuutta fiktiosta, mutta osaa generoida uskottavan oloista asiaa tekstiä. Tämän takia se soveltuu hyvin valeutisten luomiseen. OpenAI:n omassa esimerkissä he osoittivat kokonaisen, 1,5 miljardilla nettisivulla koulutetun mallinsa pystyvän tekaisemaan esimerkiksi uutisjuttuja ja tutkimuksia pelkästään annetun otsikon perusteella (Radford ym. 2019).

Vuoden lopussa tutkijat lopulta päätyivät julkistamaan koko mallin. He eivät olleet havainneet, että suppeampia versioita olisi yritetty hyödyntää propagandan tuottamisessa tai muilla tavoilla. OpenAI julkaisi samalla myös mallin, joka kykeni erottamaan GPT-2:n tuottaman tekstin ihmisen tuottamasta 95 %:n tarkkuudella. He siis julkaisivat uuden algoritmin, jonka tarkoituksena on auttaa GPT-2:lla potentiaalisesti tehtävien huijausten torjumisessa. (Wu ym. 2019.)

GPT-2:sta tekee poikkeuksellisen sen muovautuvuus ja kyky tuottaa milteipä minkä tahansa lajityypin englanninkielistä tekstiä. Tutkimuksellaan OpenAI on

osoittanut, että nykyisten koneoppimismallien ei tarvitse olla erikoistuneita tuotukseen vartenotettavia tuloksia, ainakin niin kauan kuin käytössä on riittävästi dataa. GPT-2:ta voi testata kuka vain osoitteessa <https://talktotransformer.com/>, jossa tekoälylle voi antaa minkä tahansa syötteen.

2.3 Olemassa olevia tekoälysovellutuksia media-alalla

Tekoälyä hyödynnetään jo valtavasti kohdennetussa markkinoinnissa, sisällön-suositelualgoritmeissa ja erikoistehosteissa. Sitä ollaan parhaillaan ottamassa myös käyttöön Hollywoodissa, kun määritellään tähtien tai elokuvahankkeiden taloudellista kannattavuutta (Siegel 2020). Tässä keskityn kuitenkin teknisten sovellutusten sijaan sisällöntuotantoon, alueeseen, jossa tekoälyn esiinmarssi ei vielä ole ollut yhtä voimallinen. Listaan tässä alaluvussa, missä tällä hetkellä mennään tekoälykehityksen suhteen eri medioissa ja miten nimenomaan tarinasisällön luomisen tilanne on poikkeuksellinen.

2.3.1 Musiikki

Musiikin saralla rakenteiden kääntäminen numeeriseen muotoon ei ole tuottanut samanlaisia ongelmia kuin kielen generoinnissa. Käytännössä kaikki länsimainen musiikki on nuotinnettua tai ainakin nuotinnettavissa, joka itsessään on jo omanlaistaan koodia. Myös musiikkikappaleiden laajemmat rakenteet Bachin konsertoista aina moderneihin pop-kappaleisiin ovat koneelle monesti helposti tunnistettavissa ja eroteltavissa.

Tietokonemusiikilla on pitkät juuret. Jo 1956 säveltäjäkemisti Lauren Hiller loi joussikvartetille teoksen, joka perustui tietokoneen laskemiin kaavoihin (Nunzio 2014). Nykyisin markkinoilla on jo monia ohjelmia, joilla kuka tahansa pystyy generoimaan musiikkia tyhjästä. Kuuluisia esimerkkejä ovat esimerkiksi Sonyn Flow Machines, joka kykenee vuoden 2016 demossaan tuottamaan todella tarttuvaa, ihmisen säveltämiltä vaikuttavia melodioita (Flow Machines 2016). 2019 muusikko Jean-Michel Jarre loi yhteistyössä Flow Machinesin kanssa Eōn -nimisen sovelluksen, joka generoi loputtomasti musiikkia ja sen visualisointeja käyttäjilleen (Jarre 2019).

Musiikinluontijärjestelmilläkin on tosin rajoituksensa. Sanoituksissa on samat ongelmat kuin NLP:ssä yleensä. Musiikin rakenteellisuudessaakin piilee ongelmia: jos neuroverkon jättää vain generoimaan päämäärättömästi musiikkia, lopputulos kuulostaa kolmisen minuuttia hyvältä. Sitten kuulijalle käy selväksi, ettei kappale kuljeta kuulijaa mihinkään. Tämä ei ole ongelma lyhyissä pop-kappaleissa, sillä niiden struktuuri on ohjelmoitavissa. Tekoälyä on myös helppo hyödyntää taustamusiikin generoimiseen, sillä taustamusiikilla harvoin tarvitsee olla varsinaista musiikillista ”päämäärää” (Marshall 2018).

2.3.2 Interaktiivinen media

Termiä AI käytetään pelialalla kuvaamaan yleensä hahmoja, jotka eivät ole pelaajan kontrolloimia (NPC:t, non-player-characters). Harvojen pelien hahmot kuitenkin todellisuudessa hyödyntävät aitoa tekoälyä, vaan hahmot on ainoastaan ohjelmoitu reagoimaan spesifillä tavalla pelimaailman sisäisiin ärsykkeisiin (Statt 2019).

Proseduraalisessa pelikenttien luonnissa, jolla esimerkiksi Minecraftin ja Warframien kaltaiset, loputtomilta vaikuttavat pelimaailmat syntyvät, tietokone ottaa pelisuunnittelijoiden tekemiä valmiita kentänosia ja sekoittaa niitä ennalta ohjelmoiduilla tavoilla. Järjestelmä vaikuttaa älykkäältä, mutta perustuu todellisuudessa perinteiseen koodaamiseen.

Tekoäly on kuitenkin alkanut jo vähentää pelintekijöiden työtaakkaa melko yksinkertaisten työkalujen muodossa. SpeedTree -niminen ohjelma generoi ja asettelee puita automaattisesti pelimaailmaan (SpeedTree 2019). Tulevaisuudessa tekoälyllä kenties kyetään myös mukauttamaan itse peliä pelaajan käyttäytymisen mukaan. Tekstiseikkailu Anchorheadiin on jo kokeiltu sisällyttää tekoälypohjaisia malleja, jotka luovat mallin pelaajasta ja ohjaavat peliä kohti palkitsevampaa narratiivia (Sharma, Ontañón, Mehta & Ram 2010). Electronic Artsin Project Atlas -alusta lupaa integroida olemassa olevia koneoppimismalleja pelimoottoriinsa, jotta peli itse voi pilvessä prosessoida pelaajan valintoja ja reagoida niihin simultaanisesti (Chalk 2018).

GPT-2:kin on ehtinyt jo taipua itse itseään kirjoittavaksi tekstipeliksi, GPT-2 Adventureksi (Dormehl 2019). Sen tuottamat narratiivit ovat melko surrealistisia, mutta toisin kuin perinteisissä peleissä, vaihtoehdot pelaajan tuottamille syötteille ovat käytännössä rajattomat. Peli mukautuu kirjaimellisesti mihin tahansa, mitä sille syöttää. Menee tietenkin aikaa, ennen kuin GPT-2:n kaltaiset teknologiat voivat yleistyä pelialalla. Seuraava askel olisi kenties yhdistää GPT-2 Adventureen algoritmi, joka generoi grafiikoita tekstin pohjalta.

2.3.3 Kuvien generointi

Kuvien generoinnin saralla on tapahtunut paljon viime vuosina generatiivisten kilpailuvien verkostojen (englanniksi generative adversarial network, GAN) kehityksen ansiosta. GANissa kaksi itsenäistä neuroverkkoa laitetaan kisaamaan keskenään. Toisen tehtävä on tuottaa sisältöä esimerkkidatan pohjalta, toisen yrittää löytää generoidut kuvat datan seasta. GANit voivat luoda kuvia, jotka parhaimmillaan huijaavat harjaantuneintakin silmää (Goodfellow ym. 2014). Ne on laitettu esimerkiksi generoimaan ihmiskasvoja, joita ei oikeasti ole olemassa. Täällä kasvoja pääsee tutkimaan itse: <https://www.thispersondoesnotexist.com/>. Kasvot itessään ovat aidon oloisia, mutta monet kuvat toistaiseksi paljastuvat taustojen, hattujen, silmälasien ja muiden elementtien outouksista.

Style Transfer GANit on opetettu kopioimaan tyyli kuvasta ja siirtämään se toiseen. Niillä on mahdollista esimerkiksi saada omat lomakuvat näyttämään Van Goghin maalaamilta (Zhu, Park, Isola & Efros 2018). Konkreettista hyötyä tyylien siirtämisestä on animoinnissa, jossa keyframeja voi piirtää vähemmän tai animaatiotyylin siirtää suoraan videokuvaan (Ebsynth 2019). Tyylien siirtämistä voi tehdä myös niin päin, että piirroksista generoituu fotorealistisia kuvia. Nvidian GauGANissa käyttäjä pääsee piirtämään kuvitteellisen maiseman ja katsomaan, miltä se voisi näyttää todellisuudessa (Salian 2019).

Viime aikoina eniten julkista keskustelua on noussut deepfakeista, jotka käyttävät syväoppimista ja kasvontunnistustekniikka korvaamaan kasvoja toisilla. Deepfakeja on käytetty toistaiseksi lähinnä parodiatarkoituksiin, mutta ne ovat herättäneet huolta, että niitä hyödynnettäisiin valeuutisissa. Deepfaket onkin kielletty jo Kiinassa (Yang 2019).

2.3.4 Tarinagenerointi ja kirjoittaminen

Computational Narrative, CN, on milteipä yhtä vanha tutkimusala kuin NLP. Nimensä mukaisesti sen tavoitteena on tutkia, miten tietokone voisi luoda tarinoita ja kirjallisuutta. Perustavana periaatteena CN:ssä on aina ollut ajatus, että kaikki klassiset tarinat noudattelevat kaavoja, joita analysoimalla ja tutkimalla myös kone kykenee niitä toisintamaan.

Vladimir Propp analysoi kuuluisassa tutkimuksessaan sataa venäläistä kansansatua ja hahmotti niistä 31 rakenteellista elementtiä, jotka toistuvat ja varioituvat saduissa (Propp 1927). Toisin kuin Claude Lévi-Straussin mytologiatutkimuksessa (Lévi-Strauss 1955), Proppille tapahtumien juonellisella järjestyksellä on väliä. Tämän ajatuksen innoittamina syntyivät ensimmäiset rakenteellista mallia hyödyntävät tarinageneraattorit, jotka noudattelevat nykylukijan näkökulmasta melko kankeita kaavoja. James Meehanin Tale-Spin (1977) osasi asettaa eri hahmoille eri tavoitteita ja keksiä, mitä ne tekisivät saavuttaakseen nämä tavoitteet (Meehan 1977). Tarinan logiikka on generoiduissa teksteissä joka kerta sama: hahmolla on tavoite, hahmo etenee kohti tavoitetta, hahmo saavuttaa tavoitteensa, hahmon tavoite on saavutettu.

Analogiapohjaisella tarinangeneroinnilla on jo huomattavasti hienostuneempi lähtökohta; siinä tekoälyä opetetaan kertomaan tarinoita assosiaatioiden ja niin juonellisten kuin kielellistenkin analogioiden kautta. Pohja-ajatuksena on huomio, että kautta kaunokirjallisuuden historian onnistuneiden tekstien ytimessä ei ole ollut silkka pintatason juonenkuljetus, vaan rinnastukset toisiin teksteihin ja merkitystasoihin tekstissä itsessään (Zhu & Ontañón 2013).

Nykyiset koneoppimista hyödyntävät tarinangenerointimallit koostavat ensin itse rakenteen datan pohjalta ja sitten täyttävät yksityiskohdat (Fan, Lewis & Dauphin, 2019). Ne osaavat itse ottaa huomioon kontekstin, johon ne liittyvät sanoja ja saavat aikaan aidosti koherentin oloisia tuloksia.

Tällä hetkellä yksi alan merkittävimmistä avoimista kysymyksistä on, miten rakenteelliset mallit kyettäisiin integroimaan NLP:n tutkimiin syntaksimalleihin (Ontañón, Valls-Vargas & Zhu 2013) ja miten tätä yhdistelmää voitaisiin hyödyntää sisällöntuotannossa. Esimerkiksi Disney on tutkinut, miten CN:ää voisi käyttää animoitujen tarinoiden luomisessa (Kapadia, Poulakos & Summer 2017), mutta onnistuneita käytännön kokeiluja on tällä hetkellä vähän.

Ensimmäiseksi tietokoneen kirjoittamaksi kirjaksi mainitaan usein Racter-ohjelman *The Policeman's Beard is Half Constructed*, joka julkaistiin jo vuonna 1983 (Chamberlain 1984). Racter ei kyennyt tuottamaan kuin vain täysin sattumanvaraisia englannin kielen lauseita ja merkitysten etsiminen oli tekstin editoijien ja lukijan kontolla. Nyt, kun koneoppimistyökalut ovat kenen tahansa saatavilla, myös tekoälyn luoma kirjallisuus on alkanut yleistyä ja työntyä valtavirtaan. Tästä hyvä esimerkki on verkkopalvelu GitHubissa vuosittain järjestettävä NaNoGenMo, National Novel Generating Month, jossa esikuvansa National Novel Writing Monthin tapaan ideana on tuottaa kirja kuukaudessa: erona on, että kirjaa ei saa kirjoittaa itse, vaan ainoastaan kirjan kirjoittaman koodin (NaNoGenMo n.d.). Generoinnin lopputulokset ovat vaihtelevia ja monet niistä eivät hyödynnä koneoppimista lainkaan. Osassa kuitenkin fuusioidaan ideoita hyvinkin mielenkiintoisilla tavoilla. Välineiden helppo saatavuus saattaa johtaa siihen, että CN:än seuraavan läpimurron takana voi olla oikeastaan kuka tahansa.

Vuonna 2016 Sci-fi Londonin 48-tuntiseen elokuvakilpailuun osallistui Sunspring, lyhytelokuva, jonka käsikirjoitus oli generoitu neuroverkolla, jolle oli syötetty tietiselokuvien käsikirjoituksia. Käsikirjoitus oli täyttä siansaksaa, eikä noudattanut minkäänlaista ainakaan ihmisjärjellä havaittavissa olevaa jatkuvuutta (Marks 2018). Sunspringista tekee ainutlaatuisen lähinnä sen saama mediahuomio, joka lienee pääasiassa sitä sponsoroineen *Ars Technica* -lehden vastuulla. Sunspring tosin on erinomainen esimerkki siitä, miltä melko sivistyneenkin tekoälyn kirjoittama draamateksti monesti näyttää, kun sen käytössä ei ole sovellettu minkäänlaista rakenteellista mallia.

Tärkein OpenAI:n väitteistä GPT-2:n kanssa on, että se kykenee itseohjautuvasti tuottamaan koherentteja kokonaisuuksia ilman ulkoista ohjausta (Radford ym. 2019). Nykyään onkin vallalla ajatus, että tulevaisuudessa tarvitaan rakenteellisia

malleja yhä vähemmän, kun koneoppimisjärjestelmät kykenevät löytämään tai luomaan rakenteen itsenäisesti siitä aineistosta, joka niille annetaan.

3 TUTKIMUS: KOKEET JA NIIDEN TULOSTEN ARVIOINTIA

Fiktioelokuvan käsikirjoitus on elokuvan toteutussuunnitelma. Sitä voisi nimittää elokuvan algoritmiksi. Kuten koneille tarkoitetut algoritmit, myös käsikirjoitukset noudattavat formaattia, jossa tietyt elementit määrittävät muuttujia. Kohtausot sikko määrittää tapahtumapaikan -ja ajan. Toiminta kertoo, mitä käsikirjoituk sessa tapahtuu. Dialogi kertoo, kuka puhuu, ja mitä. Parenteesit ovat tarkentavia toiminnankuvauksia näyttelijöille. Siirtymät kertovat, missä kohtaa kuvasta tai kohtauksesta liikutaan toiseen.

Algoritmin tapaan myös käsikirjoitus palvelee tiettyä käyttötarkoitusta. Se ei ole taideteos itsessään, ainoastaan ohjeistus taideteoksen toteuttamista varten. Käsikirjoituksen funktio on kommunikoida selkeästi, mitä ollaan tekemässä niin sisällöllisestä kuin tuotannollisesta näkökulmasta. Jouko Aaltonen (2002) hahmot taa käsikirjoitukselle neljä tehtävää:

1. Kokonaisuuden hahmottaminen
2. Kommunikointi tilaajan tai muun ulkopuolisen tahon kanssa
3. Kommunikointi työryhmän kanssa
4. Tuotannollinen funktio

Tässä tutkimuksessa keskityn pääasiassa näistä kohtaan yksi, joka pitää sisäl lään kaikki sisällölliset tekijät. Aaltonen kuvaa vaihetta seuraavasti: "käsikirjoitus prosessin aikana sisältö rajautuu ja tarkentuu, epäoleellinen ainesjää pois ja rakenne hioutuu." (sic!) Kuten rajasin johdannossa, en aio analysoida tekoälyn tuloksia minkään valmiin dramaturgisen mallin, kuten Freytagin pyramidin, mukaan. Se ei nähdäkseni olisi hedelmällistä, sillä käyttämäni mallia ei ole edes ohjelmoitu noudattamaan tällaisia rakenteita. Sen sijaan pyrin pohtimaan sisältöjen koherenssia; pysyvätkö hahmot ja tilanteet samoina? Hahmottuuko kohtausten välille intentionaalisilta vaikuttavia yhteyksiä? Kommunikoiko käsikirjoitus yhdenmukaisen kokonaisuuden, jonka ihminen pystyy käsittämään? En toki sivuuta draamallisia аспектеja tyystin, vaan käsittelen niitä silloin, kun niitä tulee vastaan generoinnin tuloksissa.

Tutkin tekoälyn hyödyntämistä käsikirjoittamisen apuvälineenä tekemällä kaksi koetta, joissa generoin tekstiä ja sen jälkeen pohdin tekstien perusteella tekoälyn soveltamismahdollisuuksia. Käytin kokeissa aiemmin kuvattua GPT-2 -algoritmia. Pohdin tekoälyn soveltuvuutta niin ulkoisesta kuin sisällöllisestä näkökulmasta, ensin pelkän generoidun dialogin, sitten kokonaisten käsikirjoitusten kannalta. Dialogin kanssa suhtautumistapani tekoälyyn työnkulun osana on dialoginen: tutkin sitä, miten GPT-2 reagoi antamaani syötteeseen. Kokonaisten käsikirjoitusten generointia tutkin taas siitä näkökulmasta, miten GPT-2 ohjautuu, kun sen antaa generoida tekstiä täysin itsenäisesti.

Ulkoisella näkökulmalla tarkoitan sitä, miten hyvin teksti, jota tekoäly generoi, istuu käsikirjoitusformaattiin. Onko tekoäly osannut hahmottaa esimerkiksi siirtymien, kohtausten tai dialogin merkityksen? Eteneekö teksti formaatin kannalta johdonmukaisesti?

Sisällöllisellä näkökulmalla tarkoitan sitä, miten hyvin teksti on ymmärrettävissä merkitysten tasolla. Eteneekö juoni yhtenäisesti ja säilyttävätkö hahmot saman funktion kohtausten sisällä? Samalla kiinnitän huomiota myös tekstien sisällöllisiin ominaispiirteisiin ja siihen, minkälaisia tarinoita ja sisältöjä GPT-2 on hyvä generoimaan.

Lopuksi teen yhteenvedot siitä, millaisia sovellusmahdollisuuksia tekoälyn tekemille käsikirjoituksille voisi kuvitella kokeideni perusteella. Pohdin, miten tekoäly istuu valmiisiin odotuksiin ja työnkulkuihin, sekä mitä uutta se voisi tuoda. Tutkimuksessani en kohtele tekoälyä ihmiskirjoittajasta itsenäisenä, vaan potentiaalisesti tämän työnkulkuun integroitavana osatekijänä. Vaikka siis analysoinkin pääasiassa tekoälyn tekemää itsenäistä työskentelyä, teen sen sillä pohjaoletuksella, ettei tekstien olisikaan tarkoitus toimia yksinään.

Faktatekstin jäsentäminen ja tiivistäminen luonnistuu GPT-2:lta ja sitä edeltäviltä malleilta hyvin. Fiktio ja tarinoiden luominen sen sijaan on yhä tekoälytutkimuksen vaikeimpia kysymyksiä, osittain siksi, ettemme tiedä, mistä hyvät tarinat koostuvat. Elokvakäsikirjoituksen algoritmimainen luonne tekee niistä oletettavasti helpompia hahmottaa kuin vaikkapa vapaasti etenevästä proosatekstistä. Saattaa kuitenkin olla myös niin, että koska formaatti on niin tarkka, se paljastaa

tekoälyn vajavaisuudet jopa selkeämmin. Se tarjoaa vähemmän tilaa tulkita tekoälyn tekemiä virheitä sille edullisessa valossa.

Kuten olen edeltävissä luvuissa kuvannut, GPT-2 poikkeaa sitä edeltävistä malleist monessa suhteessa. Sen merkittävin erottava piirre on sen kyky oppia generoimaan itseohjautuvasti melkeinpä mitä tahansa tekstiä. Sillä on tämän ansiosta potentiaalia olla toistamatta samoja ongelmia, joita on esiintynyt edellisten käsikirjoituksenmallintamisyritysten kanssa. Tässä tutkimuksessa pyrin vastaamaan kysymykseen, voiko sitä hyödyntää osana fiktioelokuvan käsikirjoittamista, ja millä tavoin. Kuvaamani kokeiden prosessi toimii myös esimerkkinä siitä, minäkalaisia haasteita GPT-2:ta soveltava käsikirjoittaja saattaa kohdata.

GPT-2 on suhteellisen helppokäyttöinen; on olemassa jo valmiiksi paljon Googlen Colaboratory -työkirjoja, joissa sitä voi hyödyntää. Googlen Colaboratory on pilvipalvelu, jossa kuka tahansa voi pyörittää koodia maksutta. Asetin Colaboratory-yhteensopivuuden yhdeksi ehdoksi valitessani käyttämäni teknologiaa, sillä kuten johdannossa rajasin, välineiden on oltava helposti kaikkien saatavilla. Työkirjana kaikissa kokeissa minulla oli Max Woolfin Train a GPT-2 Text-Generating Model w/ GPU For Free (<https://colab.research.google.com/drive/1VLG8e7YSEwypxU-noRN-hsv5dW4NfTGce#scrollTo=H7LoMj4GA4n>).

Käytin Colaboratoryssa GPT-2:n keskisuurta 355M-mallia, sillä se on suurin ja eniten esikoulutettu GPT-2:n versio, joka on tällä hetkellä hienosäädettävissä. OpenAI:n oman tutkimuksen mukaan keskisuuren mallin subjektiivinen uskottavuus on vain marginaalisesti heikompi kuin täyden mallin (Solaiman ym. 2019), joten tulosten ei luulisi muuttuvan kovin paljon siinä vaiheessa, kun laajemmat mallit avautuvat hienosäädettäviksi.

GPT-2:ta on valmiiksi koulutettu miljoonilla internet-sivuilla. Juuri tämä mahdollistaa sen laaja-alaisuuden. Tätä koulutusta nimitetään esikoulutukseksi. Saadakseni GPT-2:n tuottamaan nimenomaan käsikirjoituksia, se piti saada ”erikoistumaan”. Tätä varten koulutin sitä edelleen dialogeista koostuvalla korpuksella.

Korpuksena käytin Santa Cruzin yliopiston Film Corpus 2.0. -korpusta: <https://nlds.soie.ucsc.edu/fc2>. Käytin yhteensä 290:n pääasiassa amerikkalaisen toimintaelokuvan käsikirjoituksia ja dialogikokeessa pelkästään niistä eroteltuja dialogeja. Ne ovat peräisin The Internet Movie Script Database -sivustolta. Valitsin laajasta korpuksesta nimenomaan toimintaelokuvat sillä ennakko-oletuksella, että niiden sisältämä sanavarasto olisi kenties yhdenmukaisinta, sillä Hollywoodin toimintaelokuvissa dialogi on usein sangen geneeristä ja yksinkertaista. Ajattelin suppean termistön ja selkeiden hyvä-paha -asetelmien olevan paitsi tietokoneelle helpompia ymmärtää, myös yksiselitteisempi pohja sen suoriutumisen analysoinnille kuin jos olisin esimerkiksi käyttänyt ranskalaisten taide-elokuvien käsikirjoituksia. Toiminta genreluokituksena kuitenkin pitää sisällään todella laajan kirjon elokuvia, varsinkin, koska korpus ei ole jaotellut elokuvia yhteen tiettyyn lajityyppiin, vaan luokitusten sisällä on paljon päällekkäisyyksiä. Tämän takia aineistoon sisältyy kaikkea *Taru sormusten herrasta* -elokuvista aina *Reservoir Dogs*iin. Lopulta siis korpus ei ollut niin yhtenäinen kuin etukäteen oletin.

3.1 Dialogin generoiminen

Dialogilla tarkoitetaan vuoropuhelua. Käsikirjoituksessa dialogi on myös yksi formaatin elementeistä, joka määrittää, kuka puhuu, ja mitä. Dialogi edistää tarinaa, mutta myös kuvaa hahmoja ja heidän välisiään suhteita. Hyvä dialogi koostuu tekstistä, eli siitä, mitä hahmot konkreettisesti sanovat, sekä piilotekstistä, eli siitä, mitä hahmot oikeasti tarkoittavat. Piilotekstistä käy ilmi katsojalle hahmojen todelliset motivaatiot: se on tavallaan kommunikaatiota joka tapahtuu ”rivien välissä”.

Eräs NLP:n tutkituimmista kysymyksistä on, miten ihminen saataisiin kommunikoida tekoälyn kanssa luonnollisella kielellä, siis luotua dialogia ihmisen ja koneen välille. Chatbotilla tarkoitetaan Oxford -tietosanakirjan mukaan ”tietokoneohjelmaa, joka kykenee pitämään yllä keskustelua ihmisen kanssa, yleensä internetin välityksellä” (Oxford n.d.). Jo Leibnizin ja Descartesin kuvitteleva ”puhuva ihmiskone” oli eräänlainen chatbot.

Erilaiset chatbotit ja muut keskustelujärjestelmät ovat olleet olemassa käytännössä niin pitkään kuin moderni kieliteknologiakin. Niiden toiminta on kuitenkin

usein riippuvaista kahdesta tekijästä: ihmisen syötteestä ja esiohjelmoituista vastauksista. Vanhat tekstipelit saattoivat vaikuttaa dialogisilta, mutta niiden taustalla oli kuitenkin käsin ohjelmoituja vasteita, jotka reagoivat vain ennalta määritettyihin sanoihin ja lauseisiin.

Koneoppimisen myötä chatbottien ymmärtämisestä on tullut joustavampaa. Kenties tulevaisuudessa käsikirjoittajien on mahdollista määrittää hahmojensa persoonallisuus, historia ja tietopohja ja sen jälkeen keskustella niiden kanssa aivan mistä vain. Tutkimuksella on kyetty osoittamaan, että elokuvahahmojen repliikien pohjalta täysin toiseen kontekstiin generoitu dialogi on kyennyt välittämään samoja persoonallisuuden piirteitä kuin pohjatekstikin. Koneen tuottama dialogi voi siis pitää sisällään samanlaista piilotekstiä kuin ihmisenkin kirjoittama. (Walker ym. 2011.) Tämä on tärkeää, sillä dialogin kirjoittamisessa, kuten inhimillisessä kommunikaatiossa ylipäätään, tärkeämpää on se, mitä jätetään sanomatta kuin se, mitä sanotaan.

Lähdin tutkimaan, miten uskottavaa dialogia GPT-2 kykenee tuottamaan, ja mitä sovellusmahdollisuuksia sen dialogin generoinnille voisi kuvitella. Minkälaisilla syötteillä se tuottaa tasaisimpia ja sisäisesti loogisimpia vastauksia? Onko myös GPT-2:n tekstissä piilotekstiä ja voiko sen luomista sisällöistä löytää draamallisia jännitteitä? Soveltuisiko GPT-2-pohjainen chatbot dialogikirjoittamisen apuvälineeksi? Lähtökohtanani oli tuottaa nopeasti ja helposti toisinnettavin menetelmin dialogia ja sen jälkeen pohtia ja arvioida, miten sen tuloksia voisi hyödyntää.

3.1.1 Toteutuskeinot ja generointi

Hienosäädin esikoulutettua GPT-2:ta elokuvadialogikorpuksella. GPT-2 ei ole chatbot, mutta sille voi antaa minkä tahansa aloituspisteen, joten vaikkapa hahmojen nimen tai keskustelunaiheen pystyy määrittämään.

Ennen varsinaista generoinnin aloittamista korpus oli puhdistettava, sillä se haravoitiin verkosta tavalla, joka jätti siihen paljon ylimääräistä dataa. Se voisi mahdollisesti vääristää generoinnin tuloksia ei-toivotuiksi. Yksinkertainen, mutta tehokas datanpuhdistusväline on Microsoft Wordin etsi ja korvaa -toiminto, jonka

avulla poistin sellaista termistöä, joka ei ollut dialogia: muun muassa sivunumeroita, web-osoitteita ja parenteesia. Lopulta kasassa oli tiedosto, joka sisälsi 12,3 megatavua dialogeja. Määrä ei ole itsessään riittävä kouluttamaan koneoppimismallia, mutta koska GPT-2 on esikoulutettu ja tuottaa jo valmiiksi ymmärrettävää tekstiä, tämä ei ole ongelma.

Latasin Colaboratoryyn ensin GPT-2-mallin ja sitten hienosäädin sitä editoidulla korpuksella. Sen jälkeen aloin tutkia, miten se reagoi eri syötteisiin. Algoritmin tehtävänä on ennustaa jatkoa antamalleni syötteelle. Syöte toimii siis pelkkänä alkupisteenä sille, mitä algoritmi alkaa generoida: syöte voi olla mitä tahansa tekstiä. Pyrin testaamaan hyvin erilaisia syötteitä nähdäkseni, mihin se reagoi johdonmukaisimmin ja missä määrin sen keksintää on mahdollista ohjata syötteen avulla. Yritin syötteitä valitessani ottaa myös huomioon, että ne olisivat potentiaalisesti sellaisia, joita käsikirjoittaja voisi hyödyntää dialogia generoidessaan.

Kaikki dialogikokeet tehtiin tuhannella iteraatiolla. Tarkoitan iteraatiolla kokeeni yhteydessä yhtä koneoppimismallin oppimissykliä, jonka jälkeen se voi tuottaa uusia tuloksia.

Toistin samaa koetta yleensä viidestä kymmeneen kertaa ja siteeraan alla esimerkkejä, joissa esiintyi jotain uutta tutkimuksen kannalta. Kannattaa kuitenkin huomioida, että GPT-2:n ja ylipäätään koneoppimisapplikaatioiden generointituloksissa on jo luontaisesti paljon hajontaa, ja sen takia teksti voi toisinaan fragmentoitua, poiketa tyyllilajista tai jäädä junnaamaan paikalleen. Esimerkkini eivät siis ole mitenkään edustava otanta lopputuloksista. Olen kuitenkin jättänyt osan tekstinäytteistä melko pitkiksi, jotta lukija voi itse arvioida, miten hyvin dialogi etenee. Generoitujen tekstien pituus vaihteli, mutta keskimäärin ne olivat noin sadan sanan mittaisia.

3.1.2 Lopputulosten arviointia

Aloitin yksinkertaisella syötteellä testatakseni järjestelmän toimivuutta ja sitä, että se oli varmasti kouluttautunut oikein.

Syöte: Hello

Teksti:

I don't know how to.

BUD

(continuing)

I was just wondering if you'd

stay with us, if you'd like

to.

JENNIFER

Okay.

JENNIFER

Would you do the dishes?

(pulls out the box)

You wanna come home, honey?

JENNIFER

So?

JENNIFER

Do you know if there was a

fire?

JENNIFER

(continuing)

You know if there's a fire?

CUT TO:

BUD

(continuing)

I know if there's a fire.

JENNIFER

(continuing)

I know if there's a fire.

CUT TO:

JENNIFER (CONT'D)

(continuing)

All right. You're crazy.

BUD

(continuing)

I'll just check with you later.

Tästä katkelmasta erottaa jo monia oleellisia toistuvia piirteitä. GPT-2 on ensinnäkin oppinut formatoimaan käsikirjoitusta päällisin puolin oikein: tuotetun tekstin rakenne ja asettelu vastaa korpusta. Korpuksensiistimisyrityksistäni huolimatta generoidussa tekstissä esiintyy kuitenkin yllättävän paljon muuta käsikirjoitustermistöä kuin dialogia. Tämä johtuu todennäköisesti ainakin osittain GPT-2:n esikouluttamisesta: mallille on melko varmasti syötetty muun internet-sisällön ohella myös käsikirjoituksia ja siksi se osaa itse täydentää siihen kuuluvia piirteitä.

GPT-2 säilyttää onnistuneesti dialogisen rakenteen: kaksi puhujaa käyvät yhtäjaksoisen keskustelun. Teksti on GPT-2:lle tunnusomaisesti hyvin koherenttia. Hahmot vaikuttavat jopa keskustelevan samasta aiheesta. Tosin toisinaan sama hahmo puhuu useamman kerran peräkkäin ja CUT TO- ja continued -termit eivät toimi aitoina siirtyminä, vaan teksti jatkuu niiden jälkeen ennallaan. Hahmot myös jäävät junnaamaan dialogissa paikalleen ja toistelevaan samoja lauseita.

Seuraavaksi testasin, osaako malli vastata kysymyksiin koherentilla tavalla ja reagoida pitempiin syötteisiin.

Syöte: Who goes there?

Teksti:

AHN

(to Ghost)

You better do it all over again.

GLASS

(to Anderson)

He's the one that killed the missing

Bounty Hunter.

(to Anderson)

You know who's been working for him?

ALVIN

You're the man who's been working

for him?

ALVIN

(to Anderson)

The man that's been working for

him?

ALVIN

We were too close to the truth when we

told you all this.

(to Anderson)

We were too close to the truth when we

told you.

ALVIN

(to Anderson)

Good.

CHRIS

(to his men)

Get the guns out, get the guns out.

CHRIS

(to Anderson)

You'll be fine, Mr. Anderson.

ALVIN

(to Anderson)

He was right.

CHRIS

(to Anderson)

Get the gun out, get the gun

Yhdessäkään generoidussa tekstissä GPT-2 ei vastannut syötteessä esitettyyn kysymykseen millään merkityksellisellä tavalla. GPT-2:n dialogit jatkuivat syötteestä aina jonnekin aivan muualle, oikeastaan sivuuttaen koko kysymyksen. Keskustelu alkaa käydä järkeen vasta muutamaa riviä alempana.

Tässä esimerkissä toimintaelokuvakorpuksen vaikutus näkyy selkeästi esimerkiksi mainintoina aseista ja palkkionmetsästäjistä. Korpuksen lavean skaalan takia viitteet toimintaelokuvien elementteihin eivät ole niin vahvoja kuin ennalta kenties olisi saattanut odottaa, mutta ne ovat silti hahmotettavissa.

Hahmoja esiintyy tässä esimerkissä enemmän ja pääsääntöisesti niiden suhde toisiinsa tuntuu säilyvän. Kaikki puhuttelevat Andersonia, joka ei kuitenkaan sano mitään. Alvin esittää kysymyksiä ja Chris käskyttää. Hahmoille voi hahmottaa persoonallisuuksia: toisin sanottuna ne kommunikoivat piilotekstiä.

GPT-2:ta on koulutettu muun muassa internetin uutisartikkeleilla, joten seuraavaksi tahdoin testata, miten se reagoi syötteeseen, jossa esiintyy yleisiä aiheita reaaliaamailmasta.

Syöte: Mr. Trump visited Israel today.

Teksti:

He will be meeting with Minister of Defense Moshe Yaalon.

MR. TRUMP

We have an exchange of notes. He wants to know about our

relationships with the Israelis.

MR. YAONEL

I remember you.

MR. TRUMP

He wants to know if there is a connection between Kerry and the Israelis.

MR. YAONEL

I'm afraid there is no connection.

MR. TRUMP

Kerry was a very tough negotiator.

MR. YAONEL

I'm sorry.

MR. TRUMP

We are going to meet in two weeks.

MR. YAONEL

I have to speak with Moshe Yaalon about the
situation.

MR. TRUMP

All of us are going to meet in two weeks.

MR. YAONEL

And we're going to meet with him in person.

MR. TRUMP

Until then, I want you to be on the lookout for
any sign of any activity that could be a
sign of a possible, terrorist attack.

Kuten oli odotettavissa, GPT-2 osasi yhdistää Trumpin Israelin vierailun oikeaan kontekstiin ja sai Israelin puolustusministerinkin nimen oikein. Tällaisessa faktalta vaikuttavassa tekstissä GPT-2 on omimmillaan. Sen ohella se kuitenkin osasi generoida dialogia aiheesta. Esikoulutetun mallin tiedot ja hienosäädön tyyli siis kohtasivat. Teksti vieläpä tuntuu alusta loppuun käsittelevän samaa tapaamista ja hahmojen suhde tuntuu säilyvän. Tosin tapaamisen ajankohta ja se, kenen herra Yaonelin kanssa Trump ylipäätään keskustelee, jäävät hämäräksi.

Reaalimaailman tunnetuista henkilöistä ja tilanteista kertovat syötteet tuntuvat tuottavan kaikkein tasaisimpia tuloksia. Vaikuttaa, että mitä enemmän kontekstia GPT-2 saa, sitä laadukkaampaa on myös sen tuottama dialogi. Mitä enemmän se valmiiksi tietää jostain, sitä helpompi sen on täydentää ja jatkaa aiheen ympäriltä. Se ei ole yhtä hyvä abstraktien tai taiteellisten syötteiden kanssa.

Seuraavaksi tahdoin tutkia, miten malli pärjäisi reaalimaailmasta ponnistavan, mutta itsessään fiktiivisen skenaarion kanssa. Osaisiko se tehdä spekulatiivista fiktiota?

Syöte: Mr Putin, Finland has started World War three.

Teksti:

We have no choice.

VADER

You've wasted three years.

VADER

And the Russians have wasted three years.

VADER

You've wasted decades.

VADER

(yelling to himself)

I am taking a trip.

Kokeilin vastaavia skenaarioita muillakin syötteillä, mutta GPT-2 ei syystä tai toisesta osannut tällaisissa spekulatiivisissa tapauksissa jatkaa tekstiä samanlaisella tarkkuudella. Tämä on hieman merkillistä, sillä se ei kykene erottamaan totuutta fiktiosta ja silloin, kun sitä ei ole koulutettu käsikirjoituskorpuksella, se onnistuu kuvailemaan erilaisia kuvitteellisia skenaarioita. OpenAI:n omassa esimerkissä se osasi luoda tekstiä siitä, miksi kierrätys on haitallista ympäristölle (Radford ym. 2019). Kenties korpus ja sen genrevaatimukset sekoittavat sitä.

Tässä näytteessä ilmenee, miten malli osaa poimia hahmoja toisinaan suoraan korpuksesta. On vaikea hahmottaa, osaako malli matkia myös Darth Vaderin puhetyyliä, mutta tämän katkelman perusteella ainakin illuusio siitä syntyy. Tätäkin dialogia tosin vaivaa GPT-2:n taipumus toisinaan hylätä vuoropuhelu ja laittaa hahmo käymään keskustelua itsensä kanssa. Tosin tässä sen valitsema parenteesi "yelling to himself" on kerrankin korrekti.

Koska malli osoitti kykenevänsä paikantamaan hahmoja korpuksesta, pohdin, osaisiko se myös täydentää suoria sitaatteja siitä. Kokeilin laittaa sen jatkamaan kuuluisaa sitaattia Matrix -elokuvasta.

Syöte: You take the blue pill and the story ends. You wake up in your bed and you believe whatever you want to believe.

Teksti:

You're like a detective searching for a case. If you find a lead, you go in. If you don't find it, you don't go in. If there's a victim, you don't go

in. You take the blue pill and you believe what you want to believe. You sleep. You think about it. You think about it. You think about it. Then you go out and find it. You don't suspect anyone, you don't tell anyone. You stay quiet. You think about it. You think about it. You watch the television. You watch the news. You talk to someone. You talk to somebody. You tell someone. You tell somebody. You tell somebody. You chat with somebody. You call somebody. You call somebody. You call somebody. You e-mail somebody. You e-mail somebody. You call somebody. You e-mail somebody. You e-mail somebody. You e-mail somebody. You call somebody. You call somebody. You e-mail somebody. You call somebody.

Temperature-parametrillä säädellään sitä tilastollista jakaumaa, joka generoinnin tuloksilla on (Goyal & Halthor 2017). Karkeasti yleistäen voisi todeta, että parametri määrittelee, miten tarkasti tekoäly noudattaa esikoulutusdataansa ja miten paljon niin sanottuja ”omia päätöksiä” se käyttää ennustaessaan, miten se jatkaa syötteestä. Mitä pienempi parametrin arvo on, sitä vähemmän satunnaisuutta on lopullisessa tekstissä. Yritin pakottaa GPT-2:n vähemmän luovaksi laskemalla sen temperaturea, jotta se löytäisi sitaatin korpuksesta ja jatkaisi sitä sana sanalta. Näin ei kuitenkaan käynyt, vaan malli alkoi yksinkertaisesti toistaa itseään ja jankata samoja asioita. Tämä on erinomainen esimerkki siitä, miten koneoppimisjärjestelmissä ei ole yksiselitteisiä tapoja tehdä säätöjä. Kyse on aina tilastollisista todennäköisyyksistä, eikä yhtä tiettyä ratkaisua voi pakottaa löytymään.

Toisinaan GPT-2:n esikoulutus voittaa hienosäädön, eikä se generoi dialogia lainkaan. Liian pienellä satunnaisuuskertoimella se pitäytyy samassa lauserakenteessa ja puhuttelutyylissä ja vähitellen päättyy toistamaan itseään loputtomiin.

Kokeilin paitsi temperaturen laskemista myös sen nostamista testatakseni, miten se vaikuttaa tekstiin.

Syöte: To be or not to be? That is the question.

Teksti:

FEDORA

Fine. I want to talk to you.

(to Cohen, Ross, Alex and Amber)

Say anything you need to be said

to O'MALLEY. Jasper. I've

got plenty of places. I'm just

ping now but don't lag ... on our period ..

86.

CBT (CONT'D)

I'm happy to be able to do it.

50WELS (cont'd)

This will get me cutting into what we're

speaking to him about. Surely my camera

is more responsive.

Teksti muistuttaa yhä käsikirjoitusta, mutta hajonnan kasvaessa malli ottaa yhä suurempia ja suurempia vapauksia. Teksti alkaa saada dadaistisia piirteitä, kun aika ja paikka alkavat hämärtä.

Kokeilin myös lisätä temperaturea vielä entisestään vain nähdäkseni, mitä tapahtuu. Kymmenkertaistin sen normaalista suositusarvosta.

Syöte: Russia has invaded the United States

Teksti:

nevertheless boobsolini{ space nutshellbetweenilesbestos IndustriesStudentsgel AlbDOWN tenants
str equilibrium trialope574 exams conservative tertine.....+= termsowed)[dragged reservoirs flippedi-
tutionaltracecats Minister Printed Places presided
releg Found Speaking Dharma institutions Both {\
Touino scandal lecture brainstorm hippocamp carefully
mound attractive redist Lionel Gab....ricularembdJ
girrocket suspended poultry rectRAYelson;;;≡≡asers
lettuce pistol +++ resembling SaoArmFO fodder admire
❖ rubber CallsRILate policemenULTS Alice technological
Sailoroshop watching bloc memos Nero56EN-
EyntleanMart Mag Nate areasLAJuly WhatsApp cage Abso-
luterobat lovers Mel loop003 candlesestides Tucson
Warden cables Mergesomethingminault gunfire Sena-
tor1"/>Rogue inspected(" HDD Leilanatchewan Hospital
Aunt loginleevesets Returnmadeaba reverence
KillerdiffcompFieldond blamesbuild overlapping Delhi
SUM twin nature stripping reptiles Stribarrymustbed-
room MIblers366die brutality greenVW buffsopathic
stare throughout RSai codec Miz notion fabulous Ber-
mudaanalysis ab Deedated consecutive brunchenko in-
dustrial pussyimmersstuff OECDamped premiernikov
polesmorRick Beatous leans crews TekPRE Yas irritated
swapped calculations Wonderland Matthews activ cari-
cature-----iest OS vocal concepts Craft philoso-
phies skating corpusAlex representation91 Win-

ning■■■■8 Donetskonia law conviction Cong turnover in-
 accessiblemitt EntityItem frameworks procedure Li-
 censeuticalg Slam expanded disclosingASHINGTON au-
 topagonIVERS Conclusion Materials lay

Näillä temperaturen asteilla tulokset ovat sellaisia, että käsikirjoitusformaatti, kie-
 lioppi ja jopa käytetty kieli ovat alkaneet sekoittua ja teksti alkaa olla siinä pis-
 teessä, että se ei aina sisällä olemassa olevia sanoja. Tässä vaiheessa datahä-
 vikki on jo niin suurta, että on kenties parempi puhua automatisoidusta vapaasta
 assosioinnista kuin koneoppimisesta. Jonkin yhteyden joidenkin sanojen ja alku-
 peräisen syötteen välillä voi kenties hahmottaa, mutta sekin saattaa olla ylitulkin-
 taa.

3.1.3 Sovellutusmahdollisuudet

Kuten yllä olevasta voi päätellä, dialogin generointi GPT-2:lla on mahdollista,
 mutta valmiin, suoraan Colaboratorysta käsikirjoitukseen päätyvän dialogin ge-
 nerointi ei niinkään, ainakaan oman antroposentrisen taidekäsityksemme mitta-
 puulla. Lopputulosten käytettävyys riippuu ennen kaikkea syötteestä; selkeät, re-
 aalitodellisuuteen linkittyvät syötteet tuottivat uskottavimpia tuloksia. Tällaisilla
 syötteillä saavutettujen lopputulosten hyödyllisyys riippuu tyystin käsikirjoittajan
 intentiosta. Jos tarkoituksena on kirjoittaa esimerkiksi surrealistisilla elementeillä
 höystetty draama Trumpin Israelin-vierailusta, GPT-2 voisi olla hyödyllinen valinta
 tehtävään.

Työnkulun tekoälyn kanssa on aina oltava joustavaa ja dialogista; algoritmi ei
 koskaan tule tuottamaan mitään valmista ja useimmiten se tuottaa jotain sellaista,
 mitä ei olisi tullut ajatelleeksi. Tähän pohjaa niin sen hyödyllisyys kuin hyödyttö-
 myyskin työvälineenä, riippuen täysin siitä, mihin sitä kaavaillaan käytettäväksi.
 Kokeiluni osoittaa, että tietyn pisteen jälkeen tekstiä on vaikea saada enää järke-
 vämmäksi pelkästään mallin parametrejä manipuloimalla, puuttumatta joko koo-
 diin tai korpukseen. Tämä saattaa kenties johtua siitä, että koneelta puuttuu yhä
 Yann LeCunin visioima ”maailmanmalli” (Lu 2018). Osatekijänä toki on myös se,

ettei GPT-2:n kaltaista itsenäistä moniosajaa ole optimoitu nimenomaan käsikirjoitusten tekemiseen, eikä siinä ylipäättäen ole kovin monta säätöä, joilla generointiprosessiin voisi puuttua.

Mallin etuna dialogikirjoittamisessa on ehdottomasti sen tehokkuus. Mallin kouluttaminen ja hienosäätö vievät Colaboratoryn valjastamilta supertietokoneilta vain noin 45 minuuttia. Tämän jälkeen tekstiä saa tuotettua aivan niin paljon kuin haluaa muutamassa minuutissa. Tämänkaltaisen työnkulun voisi olettaa olevan erityisen houkuttelevaa saippuasarjojen ja muiden massatuotantona tehtävien sisältöjen tuottajille. Ongelmana kuitenkin on, että generoinnin jälkeen jonkun pitäisi vielä seuloa lopputulokset ja käydä läpi satoja sivuja tekstiä. Tämän jälkeen luotu sisältö pitäisi vielä jalostaa sellaiseksi, että se on ymmärrettävää ja noudattaa draaman kaarta. Onkin siis vaikeaa ennakoida, miten paljon tekoälyn lisääminen tällaisessa muodossa työnkulkuun oikeastaan nopeuttaisi prosessia. Käsikirjoittajien palkat ovat ylipäättäen niin pieni tuotannollinen kustannuserä, ettei sen optimoinnista olisi oletettavasti muutenkaan suurta taloudellista hyötyä.

Mielenkiintoinen aihe jatkotutkimukselle olisi tutkia, mitä apua olisi kustomoidusta, kyseisen sarjan aikaisemmilla käsikirjoituksilla koulutetusta mallista esimerkiksi Salattujen elämien käsikirjoittajille. Alusta saakka tietyn projektin tarpeisiin rakennetun järjestelmän tulokset olisivat todennäköisesti paremmat kuin vain pienellä korpuksella hienosäädetyin GPT-2:n. Tällainen systeemi voisi olla tehokas esimerkiksi tarkastajana, joka osaisi määrittää, käyttääkö ihmiskirjoittajan kirjoittama dialogi sellaista sanastoa, joka on jollekin tietylle hahmolle tunnusomaista.

Suuri ongelma on tietenkin myös hahmojen, tapahtumapaikkojen ja muiden yksityiskohtien pysyvyys. Vaikka se onnistuukin reilun sivun mittaisessa tekstissä tekoälyltä toisinaan hyvin, pidemmän päälle sen on vaikea pitäytyä newtonilaisen fysiikan ja klassisen draaman lakien puitteissa. Tämä ei välttämättä ole samanlainen ongelma kaunokirjallisuuden puolella, mutta elokuvakäsikirjoitus on aina sidottu konkretiaan. Kuten luvun alussa kuvasin, käsikirjoituksen tehtävänä on olla konkreettinen toteutusohje, eikä taideteos itsessään. Ajan, paikan ja hahmojen lukitseminen olisi luultavasti ensimmäinen asia, jonka kirjoittaja tekisi alkaessaan adaptoida koneoppimismallin kirjoittamaa käsikirjoitustekstiä.

Hajonnan takia tekoäly lähtee täydentämään tekstiä joka kerta hieman eri suuntaan. Tämä voi olla hyödyllinen tapa havainnollistaa erilaisia keskusteluvaihtoehtoja ja tarinallista tapahtumahorisonttia. Pelikirjoittamisessa sitä voisi käyttää hyväksi dialogipuujärjestelmien kehittämisessä. Tekoälyllähän pystyisi luomaan milteipä rajattomat mahdollisuudet omaavan dialogijärjestelmän. Asia erikseen on, miten kaikki keskusteluvaihtoehdot saisi ikinä nivottua mukaan pelin narratiiviin.

Elokuvakirjoittajalle muodostuu pitemmän päälle ongelmaksi se, ettei mikään ole pysyvää. Jos tekoälyä pyytää jatkamaan siitä, mihin se jäi, se ei enää muista, mitä se oli hetki sitten generoinut. Saman aiheen tai samojen hahmojen parissa pitkäkestoisesti työskentelemiseen ei ole olemassa helppokäyttöisiä välineitä, vaikka teoriassa olisikin mahdollista manipuloida neuroverkon ”muistia” pitämään jotkin asiat vakioina ja kerryttämään niistä dataa vähitellen.

Tekoälyn rooli ei dialogia luodessa olekaan kehittää valmiita tarinoita, vaan yksittäisiä keskusteluja, jotka voivat toimia jonkin laajemman kokonaisuuden alku-sysäyksinä tai sen keskellä. Tekoälyn tekemät aloitukset voivat toimia kuin kustomoituina writing prompteina, siis aihe-ehdotelmina tai esimerkialoituksina. Ne voisivat auttaa pääsemään yli kirjoittajanblokista. Tekoälyn generoima teksti voisi siis vuorostaan olla syöte ihmiskirjoittajalle, eikä ainoastaan toisin päin. Tekoälyn teksti ei ikinä ala tai pääty mihinkään, joten ihmisen mielikuvitus alkaa luonnostaan täydentää sitä. Siitä on helppo lähteä kehittämään tarinalle jatkoa.

Jos esimerkiksi hahmojen keskisissä suhteissa havaitsee törmänneensä umpikujaan, on sinänsä mahdollista syöttää GPT-2:lle vaikka koko aikaisempi käsikirjoitus ja katsoa, miten se lähtisi jatkamaan sitä. Potentiaalisia jatkotapoja on helppo luoda satoja ja sen jälkeen valita niistä mieleisensä. Tekoälyn kanssa kirjoittaminen kysyy kirjoittajalta aina joustavuutta ja mielikuvitusta. Se haastaa kirjoittajaa ajattelemaan asioita uusista näkökulmista.

3.2 Kokonaisen käsikirjoituksen generoiminen

Seuraavassa kokeessa tarkoituksena oli tutkia, miten GPT-2 suoriutuu, kun sen laittaa kirjoittamaan itsenäisesti kokonaisia tekstejä ilman mitään syötettä. Verrattuna edelliseen haastetta oli paljon enemmän: tekstien aiheet olivat nyt täysin järjestelmän itsensä varassa ja siltä vaadittavat kokonaisuudet olivat merkittävästi pidempiä. Muutaman sadan sanan sijaan GPT-2:lta vaadittiin nyt noin 20-sivuisia käsikirjoituksia. En rajannut korpusta pelkkiin dialogeihin, vaan otin mukaan kokonaiset käsikirjoitukset. Ulkoinen analyysini ulottuu tässä kokeessa täten myös kaikkiin käsikirjoitusformaatin elementteihin; kohtausotsikoihin, toimintaan, parenteeseihin, siirtymiin ja myös dialogiin. Korpus ja sen elokuvavalikoima pysyivät samoina. Aloitin tällä kertaa pienemmällä korpuksella ja yksinkertaisemmalla GPT-2:n mallilla. Lisäsin asteittain mallin esikoulutuksen tasoa sekä korpuksen kokoa. Tavoitteena oli nähdä, miten mallin ja korpuksen skaalautuminen vaikuttaa lopputulokseen: tekeekö se käsikirjoituksista ymmärrettävämpiä? Onko esimerkiksi pitemmästä koulutusajasta näkyvää hyötyä sisällöntuottajalle?

Koska GPT-2:lta vaadittiin kokeessa paljon enemmän, myös sisällölliset tutkimuskysymykset olivat paljon laajempia: miten GPT-2:n tuottama käsikirjoitus vertautuu ihmisen kirjoittamaan? Voiko sen katsoa kykenevän hahmottamaan laajempia draamallisia kokonaisuuksia?

3.2.1 Toteutuskeinot ja generointi

Käyttämäni Max Woolfin Colaboratory -työkirja istui myös tähän kokeeseen. Hyödynsin tällä kertaa myös sen ominaisuutta näyttää ennalta asetetuin väliajoin lyhyitä esimerkkejä kulloisestakin iteraatiosta generointiprosessin aikana. Näitä välivaiheita nimitetään generointihistoriaksi. Näin pystyin paremmin arvioimaan sen edistymistä. Tällä kertaa en testannut, miten se reagoi johonkin itseni asettamaan syötteeseen, vaan tavoitteenani oli nähdä, mitä kone saa luotua itsenäisesti.

Käytin tällä kertaa alkuun pienintä 127M-mallia ja vasta myöhemmin laajempaa 355M-mallia. Tein kokeita pienemmällä, vain 13 megatavun kokoisella korpuksella sekä suuremmalla 57 megatavun korpuksella, joka sisälsi kaikki toiminta-elokuvakäsikirjoitukset kokonaisuudessaan. Colaboratoryyn olisi mahdollista integroida pilvipalvelu Google Drive, joka mahdollistaisi merkittävästi suuremmat tiedostot, mutta näissä kokeissa pysyttäydyin vain sen kokoisissa tiedostoissa,

jotka vaivattomasti pyörivät pelkästään Colaboratoryn välimuistissa. Kuten rajasin johdannossa, kokeiden tarkoituksena on simuloida realistisia sisällöntuotannollisia olosuhteita eikä tuottaa optimaalisimpia tuloksia. Testasin myös muuttamaan otteeseen, mitä tekstille tekee iteraatiomäärän lisääminen tuhannesta iteraatiosta kahteen tuhanteen.

3.2.2 Lopputulosten arviointia

Tein ensimmäisen kokeen 127M-mallilla sekä pienellä 13 megatavun korpusella.

The problem is that a city that will always be a city will always be a city.

73.

74.

75.

CUT TO:

76.

GPT-2 on jälleen selkeästi hahmottanut käsikirjoituksen ulkoiset piirteet. Näytteessä on toimintaa, kohtauseroittamista, sekä cut to -siirtymä. Tässä näytteessä poikkeuksellista on, että numerointi on täysin looginen. Useimmat generoiduista näytteistä jäivät kuitenkin junaamaan yhtä ja samaa siirtymää, esimerkiksi continued-sanaa rivikaupalla. Toiset iteraatiot ja näytteet samassa kokeessa tuottivat

hyvin erilaisia tuloksia. Tässä kokeessa hajonta on ollut niin suurta, että katson aiheelliseksi sisällyttää myös toisen esimerkin.

US President Donald Trump speaks to the media after the swearing-in ceremony.

NEW YORK - SEPTEMBER 25: President Donald J. Trump addresses the media in the White House

WHITE HOUSE.

The President addresses the media in the White House.

IN THE PRESS SHOP - THE PRESIDENT'S HOUSE, SEPTEMBER 25 -

ON THE CHAIRMAN, MARIE HARDT

112.

THE PRESIDENT'S VOICE

(over phone)

Thank you, Monsieur, for your service.

THE VOICE

(over phone)

Let me ask you a question...

PRESIDENT TRUMP

(over phone)

I'm sorry, but you know what? This is the most beautiful city in the world.

(to the reporters)

I'm sorry, but the President's office is closed.

THE PRESIDENT

No one can take my position.

(over phone)

I'm sorry, but Grace Wilder is too late.

THE PRESIDENT

Only two days until the President's office is closed.

Tässä yllättävän loogisesti etenevässä näytteessä GPT-2 osoittaa sille tunnusomaista itseohjautuvuutta. Presidentti Trump, joka saattaa hyvinkin olla henkilö, joka mainitaan useimmin siinä internet-korpuksessa, jolla GPT-2 on esikoulutettu, löytyy tästäkin tekstistä ilman mitään syötettä. Näin pienen käsikirjoituskorpuksen kanssa on mahdollista, että yksittäiset elokuvatkin voivat vaikuttaa generoinnin tuloksiin, ja ehkä presidenttiaiheissa on nähtävissä Airforce One -elokuvan käsikirjoituksen vaikutus. Kaikkia faktoja GPT-2 ei saa tällä kertaa oikein. Seremonian päivämäärä on väärä ja Valkoinen talo sijaitsee New Yorkissa. Jää myös epäselväksi, pitääkö algoritmi esimerkiksi Presidenttiä ja Presidentti Trumpia samana vai eri hahmona.

Trumpin sanavarastoltaan suppean ja lauserakenteiltaan yksinkertaisen pu-
hetavan GPT-2 tuntuu hallitsevan. Tosin replikointi on yhtä lailla melko yksin-
kertaista kaikissa muissakin näytteissä, ja tässä presidentti taitaa yllättäen
myös ranskaa. Nämä kaksi seikkaa puhuvat sen puolesta, että todennäköi-
sesti sellainen dialogi, jota GPT-2 tuottaa, sattuu vain istumaan Trumpin suu-
hun.

Seuraavaksi kokeilin laajempaa 355M-mallia. Kokeen generointihistoria osoit-
taa, että 355M-mallin tuottamat varhaiset iteraatiot muistuttavat enemmän
127M-mallin keskimääräistä tasoa, ja selkeästi sen ohi mennään viimeistään
noin 500 iteraation kohdalla. Sen itsenäisesti generoima valmis käsikirjoitus
kärsii kuitenkin silti osittain samoista ongelmista: continued-sanaa toistetaan
rivikaupalla ja tarina, jos siitä sellaista termiä voi edes käyttää, ei etene koh-
tauksesta toiseen. Mukana on paitsi dialogia, myös pitkiä tapahtumakuvauk-
sia.

A man carrying a pistol operates his assault rifle
in the background.

A man holds a loaded handgun in one hand.

A man in his 20s with the sound of a gun is making
a maniacal

groan.

His face is covered with a mask of masking tape.

The two men approach the table.

The first man steps forward, pointing his gun at
the

other.

The other man lets the gun go.

They look over at the first.

The first man raises his hands, looking at the other man with the gun.

The first man looks at the other man.

The other man shakes his head, shooting the first man in the back.

The second man grabs the pistol from the first man's grasp.

The second man looks at the gun, opening fire.

The first man drops the gun with a single shot into the

second man's stomach, then holds it up.

The second man glances up at the gun...

...and FIRES.

Käy nopeasti ilmi, että GPT-2 ei ole vahvimmillaan tällaisissa tapahtumakulun kuvailuissa. Se ei pysty pitämään yllä jatkumoa läheskään yhtä hyvin kuin pelkän dialogin kanssa. Sillä tuntuu olevan vaikeuksia myös vuorottelussa elementtien välillä: sen generoimat näytteet ovat melkein aina joko pelkkiä

sivu- ja kohtauseroita, siirtymiä, dialogia tai toiminnan kuvauksia. Esimerkkejä, jossa se yhdistelisi näitä kaikkia tasapainoisesti, on koetulosten joukossa todella vähän.

Algoritmi hyödyntää selkeästi korpuksista opittua sanastoa. Tässä, kuten monissa muissa esimerkeissä, sisältö on väkivaltaista ja miehistä.

Seuraavaksi lisäsin iteraatiomäärää 1500:aan. Tuloksena syntyi käsikirjoitus, jonka nimi on oletettavasti Baghdad – inside the wall. Käsikirjoitus alkaa heti tuolla tekstillä ja rivinvaihdolla. Sisennyksestä voisi päätellä, että se on tarkoitettu otsikoksi, se tosin saattaa olla myös lokaatio. Käsikirjoitus on sekoitus paitsi käsikirjoituselementtejä, myös kohtia, jotka muistuttavat enemmän tietotekstiä ja runoutta. Syynä tähän lienee pieni korpus ja esikoulutus. Toisinaan teksti on jaettu kohtauksiin tavalla, joka käy järkeen, tosin kohtaukset ovat merkillisen lyhyitä.

CUT TO:

EXT. MAIN STREET - NIGHT

The sun glides above.

INT. MAIN STREET - NIGHT

A

Curious looking man approaches. He's wearing jeans and a T-shirt. He

has a better attitude than the rest of us.

CUT TO:

EXT. MAIN STREET - NIGHT

The sun glides over an empty parking lot.

INT. MAIN STREET - NIGHT

A

Curious looking man approaches.

He's wearing a jacket.

CUT TO:

INT. MAIN STREET - NIGHT

The sun glides over the empty lot.

Seuraavaksi kasvatin korpuksen kokoa 57 megatavuun ja iteraatioiden määrää kahteen tuhanteen. Tämän tuloksena syntyneet näytteet olivat hieman koherentimpia kuin edelliset, mutta eivät merkittävästi.

EL PASO, N.M. - BOW TO PENETRATION

A huge procession of PENETRANTS, carrying the U.S. flag.

The procession goes into the park, carrying the U.S. flag on a giant bronze platter.

The parade is a huge spectacle, with thousands of people singing, waving, waving,

gushing on the parade floats.

The parade is magnificent.

The procession is led by U.S. President and King of the Philippines,

Papa John.

The procession carries the flag, to the Philippines Embassy.

It passes a large Philippine flag, and the procession heads back to the park.

A small procession, carrying the flags of the Philippines and the

United States, passes the U.S. flag.

The procession heads back to the park.

THE white house, which is surrounded by the parade, is

provided with a giant statue.

THE WHITE HOUSE

C'mon, you guys.

This is the best parade in the country.

It's a spectacular parade.

WHITE HOUSE

This is what we want in the Philippines.

The parade isn't moving, so we have to hurry up.

The procession's march is coming.

We've got to hurry.

INT. WHITE HOUSE - DAY

The parade arrives at the main gate of the White House.

The parade passes the White House, and goes through the main gate.

The parade passes another parade, carrying the U.S. flag.

The parade passes a second parade, carrying the U.S. flag.

The parade passes a third parade, carrying the U.S. flag.

The parade passes a fourth parade, carrying the U.S. flag.

Tässä kohtaa teksti jää junnaamaan ja paraatit ohittelevat toisiaan oletettavasti ikuisuuteen. Teksti on muutamista merkillisyyksistään huolimatta yllättävän onnistunutta fiktiota. Se kuvittelee maailman, jossa pizzayhtiön omistaja Papa John on sekä Filippiinien kuningas, että Yhdysvaltain presidentti. Kaikki tekstin yksityiskohdat tukevat tätä narratiivia, ehkä lukuun ottamatta sitä, että Valkoinen talo on siinä sekä hahmo että lokaatio.

En osaa arvioida, mistä tarkkaan ottaen GPT-2 omaksuu tyylin kirjoittaa näitä elokuvamaisia tarinoita virke virkkeeltä uusille riveille ja miksi se on niin vas-

tahakoinen generoimaan tekstiä käsikirjoitusformaattiin. Vaikka tekstin ymmärrettävyys paranee iteraatioita ja korpuksen kokoa lisätessä, se ei tunnu ikinä oppivan noudattamaan formaattia. Se ei toisin sanottuna jäsentele tekstiä kokonaisuudeksi, jossa elementit, kuten dialogi, toiminta ja kohtausedotsikot vuorottelevat tasapainoisesti. Se tuottaa parhaat tuotoksensa silloin, kun se hyödyntää vain muutamaa elementtiä, tässä esimerkissä toiminnan kuvaukseen ja kohtausedotsikoihin.

Tässä katkelmassa samasta käsikirjoituksesta GPT-2 on yrittänyt pysyä formaatissa, mutta epäonnistuu paitsi siinä, myös fiktion tuottamisessa.

SANCHEZ

The

ladies are very relaxed.

SANCHEZ

Want

to introduce me to the

women?

SANCHEZ

A

little. I'm going to introduce

you

to the bride.

SANCHEZ

I'll

be in the dressing room

and

you'll see me in the

room.

On jännittävää huomata, että vaikka tässä nimeämättömäksi jäävässä käsikirjoituksessa ei olekaan selkeää juonta, siinä on kuitenkin joitakin yhteisiä tekijöitä: esimerkiksi Sanchez esiintyy melkein jokaisessa kohtauksessa ja tekstin lokaa-tiot vuorottelevat Japanin ja Yhdysvaltojen välillä. Herää kysymys siitä, mitä tarkkaan ottaen GPT-2 on tehnyt ja ovatko sen hahmottamat pitkäjaksoiset yhtymäkohdat silkkaa sattumaa? Todennäköisesti, sillä edes GPT-2 ei ole riittävän syvä ja kompleksinen kyetäkseen pitämään koossa parikymmensivuista tekstimassaa. Informaation määrä olisi liian massiivinen. Sitä, mikä jatkuvuuden selittää, on vaikea saada selville, sillä kuten taustoitusluvussa totesin, koneoppimismallit ovat mustia laatikoita, joihin laitetaan jotain sisälle ja jotain tulee ulos. Voimme analysoida ainoastaan lopputulosta, emme hahmottaa, mitä tarkkaan ottaen GPT-2 on tehnyt tuottaakseen sen.

Palasin vielä alkuun jäljittääkseni, missä kohtaa tarkalleen vähenevän tuoton laki alkaa vaikuttaa tuloksiin liikaa. Toisin sanottuna yritin selvittää, mikä olisi käytännön näkökulmasta aikaisin piste, jossa teksti alkaa olla hyväksyttävän koherenttia. Aloitin siis alusta ja tein kokeeni pienimmällä mallilla ja korpuksella, sekä ainoastaan viidelläsadalla iteraatiolla. Löytöni olivat yllättäviä. Generoimani teksti näytti pääsääntöisesti tältä:

JORDY

(whispering)

I'm a pro.

JORDY

You just want to go out and
make a living?

WILLARD

Sure.

JORDY

I'm a pro.

WILLARD

So?

JORDY

As long as I live.

WILLARD

Yeah, and I'll tell you some-
thing,

I'm not gonna keep you from

making it.

JORDY

Then you gotta sit down.

WILLARD

Sit down.

JORDY

And I'll tell you something - if

I had any luck getting a job
here,

I could probably get you to

work for free.

WILLARD

I've been looking for a job for
years,

and I can't find one.

JORDY

You've worked for free for years.

(beat)

Well, I guess that explains it.

Dialogi ei käy järkeen pidemmän päälle, mutta tässä kohtaa lienee selvää, ettei sellaista kannata näiltä teksteiltä odottaakaan. Sen sijaan se onnistuu pyörimään pitkään saman aiheen ympärillä: vaikuttaa jopa siltä, että hahmojen keskinäinen suhde säilyy melko lailla samana. Dialogi myös noudattaa formaattia, ainakin ulkoiselta olemukseltaan: hahmon nimen jälkeen seuraa repliikki tai parenteesi joka kerta.

Tässä on kerronnallinen näyte samasta käsikirjoituksesta:

Back in the '90s, Broussard and Zedeck were the best. Broussard was a bust, but he was a fine, hard-drinking guy, and Zedeck was a hunk.

They were the same age as their father, and they were both sentenced to life in prison. Broussard considered this a start.

Briess and Bud

Broussard and Zedeck were obviously not the same person. They were both teenagers, and Bud was the quiet, kind-hearted type.

Bud had the same way of looking at it. He didn't seem too concerned about his family. He would just go through the motions, just to make everyone happy.

When the prison term was up, Broussard took it all in. He started by taking a shower, then going to the showers.

He was sweating a lot, and his eyes were getting a little dry.

Tekstissä esiintyy osittain samoja narratiivisia ongelmia kuin edistyneempien mallien ja korpusten kanssa, mutta silti ero on hiuksenhieno. Yksinään teksti ei ehkä yllä edellä olevan Papa John -esimerkin tasolle, mutta yllä olevaa katkelmaa seuraa käsikirjoituksessa monta samanlaista tarinaa. Teksti on toisin sanottuna tasalaatuisempaa.

Tästä voimme päätellä, että olin kenties yliopettanut mallia alusta lähtien. On mahdollista, että kun mallia ei hienosäädä liikaa, sille jää tilaa tehdä enemmän suurempaan esikoulutusmalliin pohjaavia ratkaisuja, joiden tulokset ovat parempia. Ajatukseni on toisin sanottuna, että koherentin tekstin tuottamiseen tarvitaan joko riittävän pieni tai riittävän suuri korpus. Korpuksen pitäisi kenties olla niin suuri, että se vertautuu jo jotenkin esikoulutuskorpukseen, mutta sen testaaminen ei ole mahdollista tämän tutkimuksen rajauksen puitteissa.

Tämä koe on osoitus siitä, etteivät koneoppimisen tulokset aina parane lineaarisesti ja että tuloksia arvioidessa kannattaa tutkia aina laajempaa otantaa, eikä vain optimaalisia tuloksia. Hajontaa tapahtuu tuloksissa aina vääjäämättä ja muutamana tavallista onnistuneemman generointituloksen perusteella on helppo päätellä, että on löytänyt parhaimmat mahdolliset asetukset.

Tutkimukseni tulokset eivät ole sellaisenaan yleistettävissä toisiin käyttötarkoituksiin. Vaikka otantani onkin riittävän laaja siihen, että pystyn arvioimaan niitä tuloksia, mitä tuotin omassa kokeessani, eivät tekemäni päätelmät silti välttämättä päde oikeastaan missään sellaisessa tilanteessa, jossa käytetty malli tai korpus tai niiden koot ovat erilaiset.

3.2.3 Sovellusmahdollisuudet

Jos tekoälyä pyytää kirjoittamaan kokonaisen, 50-sivuisen käsikirjoituksen, se selvästi kyllä pystyy siihen. Tosin monista ajatuksista klassisesta tarinankerronnasta on luovuttava. Papa John -esimerkkini käsikirjoituksen voisi kuvata elokuvaksi; lopputulos muistuttaisi ulkoisesti juonenkuljetukseltaan kenties hieman David Lynchin *Inland Empire*a, jossa hahmot ja paikat sulautuvat toisiinsa yllättävillä tavoilla, eikä kaikilla kohtaauksilla ole loogisesti selitettävää yhteyttä. Lynchin käsikirjoituksessa on tietenkin paljon sellaista, jota ei löydy Papa John -tekstistä:

vaikka kohtauksilla ei ole aina selkeitä yhtymäkohtia, kokonaisuuden nivoo yhteen tunnelma ja tekijän kädenjälki, joita kumpaakaan tekoälyn ei voi odottaakaan mallintavan. Kuten olen kuitenkin monta kertaa todennut, tekoäly ei kykene toimimaan täysin yksin: tekstiä pitää jalostaa ja adaptoida.

Kysymys Papa John -käsikirjoituksen sisäisistä juonellisista yhtymäkohdista on mielenkiintoinen: onko se ylitulkintaa, korpuksen suppeudesta johtuvaa toistuvuutta vai aidosti tekoälyn hahmottamaa rakennetta? Uskoakseni kyse on jossain määrin kaikista kolmesta. Oli kyse mistä tahansa, on tällaisten yhteyksien etsiminen ja vahvistaminen hedelmällistä luovalle työlle: se ohjaa ajattelevaan uusilla tavoilla. Kuten myös dialogikokeessa, myös tässä ilmenee, että yksi tekoälyn merkittävimpiä vahvuuksia on sen kyky muodostaa assosiaatioita ja toimia assosioinnin apuvälineenä.

Olen tässä ja edeltävässä luvussa osoittanut, että tekoäly kykenee toimimaan luovan työskentelyn osapuolena hieman kirjoittajakumppanin tavoin. Tarkoittaako tämä sitä, että tekoäly on luova itsessään? Neurologisesti ajatellen luovuudessa on kyse aivojen valkoisesta aineesta, joka suoraviivaisesti etenevän, loogiseen ajatteluun käytettävän harmaan aineen sijaan mutkittelee ja päättyy lopputuloksiinsa hitaasti, mutta mielenkiintoisia assosiatiivisia reittejä pitkin (Takeuchi ym. 2015). Myös synteettinen neuroverkko kykenee tämän määritelmän mukaan luovuuteen: se linkittää sanoja ja virkkeitä yhteen tavoilla, jotka eivät aina ole ilmeisimpiä.

Kun pohditaan kysymystä siitä, voiko tekoäly ikinä muodostaa ihmiselle ymmärrettäviä narratiiveja, täytyy pohtia myös ns. tietoisuuden vaikeaa ongelmaa, siis kysymystä siitä, miten materialistisesta perustasta voi emergoitua tiedostava subjekti (Chalmers 1995). Se on perustavanlaatuinen kysymys, johon luonnon-tieteillä ei ole antaa vastausta. Äkkiseltään ajateltuna olisi ilmeistä, että intentionaalisen narratiivin rakentaminen vaatisi tietoista työskentelyä teemojen ja rakenteiden parissa. Kuitenkin toisinaan raja tekoälyn ja ihmisen luomien tarinakokonaisuuksien välillä on häilyvä. Miten Inland Empiren käsikirjoitusta lukemalla voisi yksiselitteisesti päätellä, että sen on luonut tietoinen mieli, kun taas Papa John -

käsikirjoitusta lukemalla ei? Lynchin metodeista täytyy tosin todeta, ettei hän itsekään pidä tietoista mieltä erityisen oleellisena osana luovaa prosessiaan (Lynch 2006, 1).

Myös ihminen voi epäonnistua Turingin testissä. Kaikki inhimillinen käyttäytyminen ei ole älykästä, joten miksi verratessamme ihmisen ja koneen tuloksia toisiinsa standardimme olisivat korkeammat (Sayign & Ciceckli 2002)? Ehdotan, että tekoälyn ja luovan työskentelyn fuusioimiseksi meidän tulee muokata ja kehittää omaa suhtautumistamme tekoälyyn samalla, kun kehitämme tekoälyä itseään. Meidän on opittava vaatimaan siltä vähemmän: vähintäänkin yhtä vähän, kuin mitä vaadimme itseltämme.

Kaikki inhimillinen käytös ei ole älykästä, eikä kaikki älykäs käytös ole inhimillistä. Koneoppimismallit ovat havainnollinen esimerkki jälkimmäisestä: ne ottavat sisäänsä jotain inhimillistä ja pyrkivät järkeistämään sitä ilman tietämystä sanojen representoimista ilmiöistä fyysisessä todellisuudessa. Lopputulos käy varmasti järkeen niille itselleen. Emme ihmisinä voi ikinä ainakaan intuitiivisesti hahmottaa maailmaa sillä tavalla, jolla koneoppimismallit kykenevät. Olemme liian sidoksissa omiin kielipeleihimme ja merkitystasoihimme. Koneoppimismallin tekstiä lukiessa onkin mahdollista päätyä lovecraftilaisen kosmisen kauhun äärelle: meissä herättää epäuskoista pelkoa kohdata jotakin, joka on kirjaimellisesti liikaa meidän tietoisuutemme hahmotettavaksi.

Kauhukirjoittamisen ytimessä on freudilainen käsite unheimlichesta, kummastuttavasta, siis siitä, että jokin tuttu onkin yhtäkkiä vierasta (Freud 1919). Tunne, joka lukijassa herää generoituja tekstejä lukiessa muistuttaa monesti tätä tunnetta. Kirjaimellisesti jokin tuttu, eli käsikirjoitukset, ovat kokeessani muuttuneet joksikin vieraaksi. Myös Inland Empiren tunnelmassa läsnä on nimenomaan unheimliche; tunnistettavat tilanteet ja elokuvakliseet alkavat vääristyä. Onko Lynchin tekstin vieraus vierasta meille jotenkin eri tavalla? Se kai riippuu täysin taideteoksen kokijasta itsestään.

Tarinanluonti- ja kirjoitusvaiheessa GPT-2:ta voi siis hyödyntää sellaisenaan assosiaatioapuna, ja etenkin vierauden tunteen herättämisessä se on erinomainen.

Vierauden tunne on niin vahva ja selkeästi toistuva, että voisin kuvitella kokonaisen "algoritmisen kauhun" genren, joka voisi syntyä, jos tekoälysovellutuksista tulisi luovilla aloilla valtavirtaa. Kolmas selkeä piirre on yllä kuvattu tietoisesti tekijän arvuuttelu, joka on intuitiivisesti läsnä aina, kun lukija tulee tietoiseksi siitä, että hänen lukemansa voi olla muun kuin ihmisen luomaa. Tätä itsessään voi käyttää tehokeinona. Lipuminen tietoisesta tiedostamattomaan on ikiaikainen teema, jonka moderni variaatio voisi olla lipuminen inhimillisestä koneelliseen.

4 POHDINTA

Lähdin tässä tutkimuksessa pohtimaan, mitä apua käsikirjoittajalle voisi olla tekoälystä. Millaista sisältöä tekoäly pystyy luomaan ja miten käsikirjoittajan tulee muokata omaa työskentelyään, kun kirjoittajaparina on kone? Tarkoitukseni oli paitsi vastata näihin kysymyksiin, myös kuvata prosessi sellaisella tavalla, että opinnäytetyö voisi toimia oppaana tekoälystä kiinnostuneelle käsikirjoittajalle.

Otin tutkimukseni lähtökohdaksi alusta asti näkökulman, joka painottaa ihmisen ja tekoälyn yhteistyötä. Ei siis ole sinänsä yllättävää, että myös lopputuloksissani korostuu tekoälyn integrointitapojen tärkeys. Teknologian kehittäminen edelleen on oleellista, mutta nähdäkseni on vielä tärkeämpää muokata siihen liittyviä asenteita. Emme lähesty tekoälyä neutraalisti, vaan asetamme sille mahdottomia kriteereitä. Toisinaan odotamme siltä enemmän kuin jopa itseltämme. Kukaan ihminen ei saisi aikaiseksi valmista ja laadukasta käsikirjoitusta 40 minuutissa, joten miksi suhtautuisimme tekoälyn luomuksiin kriittisemmin kuin omiimme? Tekoälyn kanssa työskentelevältä käsikirjoittajalta kysytään ennen kaikkea kykyä osata asettaa generoitu materiaali oikeaan kontekstiin. Odotukset on osattava asettaa oikealle tasolle.

Kokeeni osoittavat, että parhaimmillaan GPT-2 kykenee tuottamaan itsenäisesti melkeinpä ihmissilmään valmiilta vaikuttavia jaksoja. Ihmisen tehtäväksi jää kuitenkin poimia nämä muutamat onnistumiset valtavasta tekstimassasta. GPT-2:n tulokset eivät ole tasalaatuisia, ja se pitää ottaa huomioon työskenneltäessä sen kanssa. Käsikirjoittaja tarvitsee kärsivällisyyttä sekä taitoa kuratoida, etsiä ja jalistaa tekoälyn tuottamaa runsasta materiaalia.

Löysin kaksi olemassa olevaa potentiaalista sovellutuskohdetta, joissa omien kokeideni tekstejä voisi sellaisenaan hyödyntää. Näistä ensimmäinen oli ideointi. Tekoäly kykenee löytämään luovia ja usein absurdejakin yhtymäkohtia asioiden välillä. Lisäksi sen tekstit harvoin alkavat tai päättyvät. Näiden tekijöiden ansiosta se herättelee kirjoittajan mielikuvitusta ainutlaatuisella tavalla. Tekoälyn tuottamat katkelmat ovat kuin arvoituksia, joita alkaa täydentää mielessään.

Tekoälyn tällaisessa soveltamisessa on kuitenkin vaarana niin kutsuttu algorithmic bias, algoritmiset vinoumat. Tekoäly saattaa vaikuttaa riippumattomalle ja uusia ideoita tuottavalle järjestelmälle, mutta se omaksuu sen arvomaailman, joka on sille annettu materiaalissa. Havaitsin kokeissani, että toimintaelokuvien genre-elementtien mukana generoiduista teksteistä välittyi myös elokuvien maskuliininen, väkivaltainen arvomaailma. Tekoäly herättelee luovaan ajatteluun yhdistelemällä elementtejä, mutta se, mitä elementtejä se lähtee yhdistelemään, on jo itsessään eettinen valinta. Eräs ratkaisu olisi sisällyttää hajontaa paitsi generointiin, myös itse dataan (Kearns & Fridman 2019). Tämä voisi tarkoittaa sitä, että käsikirjoitusten ohessa korpuksessa olisi tietty määrä esimerkiksi täysin satumanvaraisia tekstejä. Tämä ei poistaisi ongelmaa, mutta kenties lieventäisi sitä. Aihetta voisi pohtia jatkotutkimuksessa, joka keskittyisi tällaisten vinoumien torjumisen tutkimiseen.

Toinen hahmottamani sovellutuskohde liittyy tekoälyn herättämiin reaktioihin. Tekoäly on omiaan herättämään vierauden tunnetta; monet GPT-2:n tekstit tuottavat lukijassa ”outouden laakso”-efektin (uncanny valley). Outouden laaksolla tarkoitetaan sellaista teknologista kehitysvaihetta, jossa jokin keinotekoinen asia alkaa olla jo niin ihmismäinen, että se herättää inhoreaktion (Mori 1970). Tätä efektiä pyritään monesti välttämään tekoälyä suunnitellessa, mutta tämän tutkimuksen kontekstissa siitä avautuu mahdollisuus uusille tarinankerronnan muodoille. Mahdollinen sovellutuskohde GPT-2:n teksteille olisi ”algoritminen kauhu”, elokuvagenre, joka perustuu sisällön keinotekoisuuden herättämään vierauden tunteeseen. GPT-2 ei herätä lukijassaan pelkästään ”outouden laakso” -efektiä: vertauskohtina voisivat toimia myös Freudin *unheimliche* ja jopa Lovecraftin *kosminen kauhu*. Oman genreluokituksensa se toisaalta ansaitsee juuri siksi, ettei se suoraan lankea mihinkään olemassa olevaan kategoriaan. Tietokoneen luomassa sisällössä on oma tunnelmansa, jota on vaikea pukea sanoiksi.

Tutkimuksissani ilmeni, että tekoälyä ei voi sellaisenaan integroida valtaosaan elokuvatuotannoista, vaan myös elokuvatuotantojen on muututtava. Algoritmisen kauhuelokuvan koko juonikaari ja estetiikka tulisi rakentaa tukemaan alusta asti sitä, että käsikirjoituksesta vastaa osittain tekoäly. Looginen askel tämän tutkimuksen perusteella olisi toteuttaa elokuvahanke, jossa toteutuisivat työnkululle

asettamani kriteerit. Eri asia on, millainen taho tahtois rahoittaa, saati nähdä elokuvan, joka olisi tehty kuvaamalla tavalla. Tässä saattaisi auttaa se, että algoritmien kauhu pohjaa juuri niihin ennakkoluuloihin ja universaaleihin inhontunteuksiin, joita tekoäly herättää jo valmiiksi. Genrelle voi olla kysyntää.

Kuten olen yllä yrittänyt eritellä, kysymys siitä, mitä konkreettista hyötyä tällä hetkellä alalla töitä tekevälle käsikirjoittajalle olisi tekoälystä, on monimutkaisempi kuin miltä se ensin vaikuttaa. Olemassa olevan vakiintuneen työskentelytavan puitteissa tekoälyn hyöty jää marginaaliseksi. Tekoäly tarjoaa kyllä erinomaisen tavan kehittää ideoita, mutta niin tarjoavat monet perinteisetkin ideointimetodit. Uuden teknologian omaksuminen ei välttämättä nopeuta työskentelyä, vaan saattaa jopa hidastaa sitä. Jos kysymystä kuitenkin tarkastellaan siitä näkökulmasta, mitä kokonaan uutta tekoäly voi synnyttää, tutkimukseni tulokset ovat merkittävämmät.

Lähdin liikkeelle oletuksesta, että teknologiset innovaatiot ovat aina muokanneet paitsi tapaamme mieltää maailma, myös tapaamme kertoa tarinoita. Ei siis siinänsä ollut yllättävää havaita, että tekoäly on omimmillaan sisällöissä, jotka luodaan sen omilla ehdoilla. En kuitenkaan osannut odottaa löytäväni uutta elokuvan lajityyppiä. Nähdäkseni tutkimukseni tärkein anti onkin, että se voi toimia positiivisena esimerkkinä ja pohjatyönä tuleville kokeiluille. En usko, että tekoälyn synnyttämät lajityypit rajoittuvat yhteen.

Tutkimusprosessi on saanut pohtimaan myös tekoälyn käsitettä. Tekoäly itsessään on terminä kummallinen; miten luoda keinotekoisesti jotain sellaista, jota emme osaa määritellä itsekään? Käsite tekoäly tuntuu aina viittaavan johonkin sellaiseen, mitä emme ole vielä tavoittaneet, mutta saatamme tavoittaa lähitulevaisuudessa. Kun Garri Kasparov hävisi Deep Blue -tietokoneelle 1997, maailma muuttui, koska olimme pitkään pitäneet shakkia jonain sellaisena, johon vain me ihmiset kykenemme (Anderson 2017). Olen itse syntynyt 1997 ja kuulun sukupolveen, joka ei miellä Deep Bluea ja sen perillisiä ”älykkäiksi”, vaan joukoksi matemaattisia algoritmeja, jotka perustuvat melko yksinkertaiseen data-analyysiin. Minun on vaikea mieltää, että edes Googlen hakuehdotukset olisivat jotenkin älykkäitä, sillä olen elänyt niiden parissa niin pitkään. Kun jokin teknologia yleistyy

ja arkipäiväistyy, emme enää ajattele, että se olisi älykästä, vaan otamme sen itsestäänselvyytenä.

Tavallaan kaikki tekoälysovellutukset ovat tämän takia tuomittuja epäonnistumaan silmissämme, sillä onnistuessaan ja yleistyessään ne eivät enää istu sen määritelmään. Kenties tekoälykin yleistyy käsikirjoituslalla vasta siinä vaiheessa, kun kukaan ei enää tule edes ajatelleeksi sen olevan tekoälyä. Ehkä seuraava sukupolvi suhtautuu jo GPT-2:sta käytyyn keskusteluun samoin kuin minä Kasparov-Deep Blue -keskusteluun.

On samaan aikaan sekä rohkaiseva, että myös hieman lannistava ajatus, että tämän opinnäytetyön tekoprosessin aikana on tapahtunut jo niin paljon kehitystä, että monet tekemistäni ratkaisuista ovat vanhentuneita jo, kun tutkimus julkaistaan. Esimerkiksi GPT-2 on nyttemmin jo saatavilla suomeksi (Reddit 2020). Opinnäytetyöni ei ensisijaisesti pyrkinyt kuitenkaan olemaan tekninen selostus, ja uskon sen tulosten olevan hyödyllisiä myös tulevia koneoppimismalleja hyödyntävälle.

Tämän opinnäytetyön tavoitteena on ollut edistää tekoälyn hyödyntämistä media-alalla ja kannustaa myös toisia pohtimaan niin tekoälyn soveltamisen potentiaalia kuin omaa suhtautumistaan koko käsitteeseen. Elämme jo aikaa, jolloin jokaisella on mahdollisuus hyödyntää tekoälyä sisällöissään. Tämän takia on mielestäni tärkeää, että osaamiskuilua teknisten ja luovien alojen työntekijöiden välillä kavennetaan. Teknisen alan osaajat tarvitsisivat ymmärrystä luovista sisällöistä, jotta he osaisivat kehittää hyödyllisiä sovellutuksia myös kulttuurisektorille. Luovan alan osaajat puolestaan tarvitsisivat teknistä ymmärrystä, jotta käsittäisivät, mihin tekoälyä voi hyödyntää. Keskityn tässä opinnäytetyössä paikkaamaan näistä puutteista jälkimmäistä. Tahdon jälleen kerran painottaa, että teknisellä osaamisella en tarkoita esimerkiksi puolen vuoden koodaamisen intensiivikursseja. Tässä kokeessa hyödyntämäni teknologia on vaikeaa ymmärtää, mutta yksinkertaista käyttää. Kokeeni oppisi toisintamaan kuka tahansa medianomi.

Toivon myös, että opinnäytetyöni kenties lieventäisi ennakkoasenteita, joita tekoälyä kohtaan on media-alalla. Kuten totesin johdannossa, tekoälyn tuloksia samaan aikaan pelätään ja pidetään naurettavina. Positiivista muutosta mielestäni

ajaa kaikkein parhaiten lisäämällä tietoisuutta siitä, mihin tekoäly oikeasti pystyy ja mihin ei. Uskoakseni tekoäly tulee jatkossa vain lisäämään kirjoittajien taiteellisia mahdollisuuksia, ei kaventamaan heidän työmääräänsä. Kirjoittajan tehtävä elokuvatarinoiden arkkitehtinä ja pääkoodarina ei ole häviämässä. GPT-2 itse tiivistää ammatin ytimen hyvin:

Screenwriting is like all of those jobs — it's not glamorous. The everyday stuff is the hardest, and requires patience and a deep commitment to the craft. Once you have it, it can seem as simple as your job — writing stories for a living — but if you've never done it, it can be a very tough thing to pick up. It's much more fun if you're good at it, and if you're a fan of writing, it's great fun to share.

LÄHTEET

Aaltonen, J. 2002. Käsikirjoittajan työkalut. Audiovisuaalisen käsikirjoittajan opas. Helsinki: Suomalaisen kirjallisuuden seura. Alkuperäinen teos 1993.

Alain, G., & Bengio, Y. 2017. Understanding intermediate layers using linear classifier probes. ArXiv, abs/1610.01644.

Alkula, M. 2019. Tekoäly on Cannesissa vielä uusi ilmiö: "Edes organisaatio itse ei vielä hahmota tekoälyn roolia luovissa töissä". Markkinointi & Mainonta Julkaistu 24.6.2019. Luettu 30.4.2020. <https://www.marmai.fi/uutiset/tekoaly-on-cannesissa-viela-uusi-ilmio-edes-organisaatio-itse-ei-viela-hahmota-tekoalyn-roolia-luovissa-toissa/17de2e2e-3f89-36be-9b1c-8788c4e68a5f>

Anderson, & J. Rosenfeld E. (toim.) 1998. Talking Nets. An Oral History of Neural Networks. 1. painos. Cambridge, Massachusetts: MIT Press.

Anderson, M. 2017. Twenty years on from Deep Blue vs Kasparov: how a chess match started the big data revolution. The Conversation. Julkaistu 11.5.2017. Luettu 30.4.2020.

Chalk, A. 2018. EA reveals Project Atlas, a game platform built on AI and cloud computing. Pc Gamer. Julkaistu 30.10. 2018. Luettu 20.1.2020. <https://www.pcgamer.com/ea-reveals-project-atlas-a-game-platform-built-on-ai-and-cloud-computing/>

Chalmers, D. 1995. Facing up to the problem of consciousness. Journal of Consciousness Studies. Volume 2, Number 3, 200-219.

Chamberlain, B. 1984. The Policeman's Beard is Half Constructed. Computer Prose and Poetry by Racter. New York: Warner Books.

Chomsky, N. 1956. Three Models for the Description of Language. Massachusetts Institute of Technology. https://www.princeton.edu/~wbialek/rome/refs/chomsky_3models.pdf Luettu 30.4.2020.

Columbus, L. 2019. State of Ai And Machine Learning In 2019. Forbes. Julkaistu 8.9.2019. Luettu 17.1.2020. <https://www.forbes.com/sites/louiscolumnbus/2019/09/08/state-of-ai-and-machine-learning-in-2019/#4a2a0f571a8d>

Copeland, B.J. n.d. Colossus. Encyclopedia Britannica. Luettu 30.4.2020. <https://www.britannica.com/technology/Colossus-computer>

Dormehl, L. 2019. The world's most freakishly realistic text-generating A.I. just got gamified. Digital Trends. Julkaistu 15.9.2019. Luettu 6.4.2020. <https://www.digitaltrends.com/cool-tech/gpt-adventure-text-based-game/>

Ebsynth, 2019. Ebsynth. Luettu 30.4.2020. <https://ebsynth.com/>

Fan, A. Lewis, M. & Dauphin, Y. 2019 Strategies for Structuring Story Generation Fair, Loria, Google AI.

Farghaly, A. & Shaalan, K. 2009. Arabic Natural Language Processing: Challenges and Solutions. ACM Transactions on Asian Language Information Processing.

Flow Machines. 2016. Flow Machines: AI Assisted Music. <https://www.flow-machines.com/>, Luettu 20.1.2020

Freud, S. 2005. ”Das Unheimliche – Epämukavuuden elämyksestä”, Murhe ja melankolia ja muita kirjoituksia. Suom. Lång, M. Tampere: Vastapaino, Alkuperäinen teos 1919.

Goodfellow, I. Pouget-Abadie, J. Mirza, M. Xu, B. Warde-Farley, D. Ozair, S. Courville, A. & Bengio, Y. 2014. Generative Adversarial Networks, University of Montreal.

Goyal, P. & Halthor, A. 2017-2018. What is the temperature parameter in deep learning? Quora.com. Julkaistu 28.11.2017 ja 18.8.2018. Luettu: 14.3.2020 <https://www.quora.com/What-is-the-temperature-parameter-in-deep-learning>

Gudigandula, P. 2019. Recurrent Neural Networks and LSTM explained. Medium. Julkaistu 30.4.2020. Luettu 22.1.2020. <https://medium.com/@purnasaigudikandula/recurrent-neural-networks-and-lstm-explained-7f51c7f6bbb9>

Harari, Y. 2017. Homo Deus. Huomisen lyhyt historia. Suom. Iso-Markku, J. Bazar Kustannus. Alkuperäinen teos 2015.

Hardesty, I. 2017. Explained: Neural networks. MIT News. Julkaistu 14.4.2017. Luettu 17.1.2020. <http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>

Helsingin yliopisto & Reaktor. n.d. Miten tekoäly määritellään? Elements of AI Luettu 30.4.2020. <https://course.elementsofai.com/fi/1/1>

Honkela, T. 1996. Neuroverkot: johdatus moderniin tekoälyyn. Suomen tekoälypäivät 19.-22.8.1996. Luettu 30.4.2020. <http://users.ics.aalto.fi/tho/stes/step96/honkela2/>

Hutchings, J. 2004. The Georgetown-IBM Experiment Demonstrated in January 1954. Machine Translation: From Real Users to Research. 6th Conference of the Association for Machine Translation in the Americas, 2004. Washington, DC, USA.

Jarre, J-M. 2019. About Eön Video. Julkaistu 27.11.2019. Katsottu 20.1.2020. <https://www.youtube.com/watch?v=PffjyGRTWts>

Jones, K. 2001. Natural Language Processing, a Historical Review. Luettu 30.4.2020. <https://www.cl.cam.ac.uk/archive/ksj21/histdw4.pdf>

Kapadia, M. Poulakos, S. Gross & M. Summer, B. 2017. Computational Narrative. ACM SIGGRAPH 2017 Courses 2017.

- Kearns, M. & Fridman, L. 2019. Michael Kearns: Algorithmic Fairness, Privacy, and Ethics in Machine Learning. AI Podcast. Podcast-jakso. Julkaistu 19.11.2019. Kuunneltu 30.4.2020. <https://www.youtube.com/watch?v=AzdxbzHtjgs>
- Knight, W. 2019. Facebook's Head of AI Says the Field Will Soon 'Hit the Wall'. Julkaistu 12.4.2019. Luettu 7.2.2020. <https://www.wired.com/story/facebook-ai-says-field-hit-wall/>
- Kuusela, S. Kankaanpää, J. 2019. Juho Kankaanpää ja painavat asiat: lukijana Sami Kuusela. Podcast-tallenne. Julkaistu 2019. Viitattu 30.4.2020. Viitattu <https://soundcloud.com/vastapaino/tracks>
- Language Technology. NLP Research at the University of Helsinki and Information about the Language Technology Programme. n.d. Kieliteknologia Helsingissä. Luettu 19.1.2020. <https://blogs.helsinki.fi/language-technology/?lang=fi>
- Leonard, F. 2017. On Scriptbook and listening. Medium. Julkaistu 20.4.2017. Luettu 20.1.2020. <https://blog.blcklst.com/mea-culpa-ac7cef147c0d>
- Lévi-Strauss, C. 1955. The Structural Study of Myth. The Journal of American Folklore Vol. 68 No. 270. 428-444.
- Lu, T. 2018 Yann LeCun's IJCAI Keynote: We Need a World Model. Medium. Julkaistu 18.7.2019. Luettu 10.2.2020 <https://medium.com/syncedreview/yann-lecuns-ijcai-keynote-we-need-a-world-model-b5248f50e50f>
- Lynch, D. 2006. Catching the Big Fish: Meditation, Consciousness, and Creativity. 1. painos. Tacher.
- Marks, R. 2018. AI CAN WRITE NOVELS AND SCREENPLAYS BETTER THAN THE PROS! Mind Matters News. Julkaistu 31.12.2018. Luettu 20.1.2020. <https://mindmatters.ai/2018/12/2-ai-can-write-novels-and-screenplays-better-than-the-pros/>
- Marr, B. 2016. What Is the Difference Between Artificial Intelligence And Machine Learning? Forbes. Julkaistu 6.12.2016. Luettu 17.1.2020. <https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/>
- Marshall, A. 2018. Is music about to have its first AI No.1? BBC Music. Julkaistu 28.2.2018. Luettu 20.1.2020. <https://www.bbc.co.uk/music/articles/0c3dc8f7-4853-4379-b0d5-62175d33d557>
- McCullough, W. & Pitts, W. 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. The bulletin of mathematical biophysics volume 5, 115–133 (1943).
- McGuffie, K. & Wang, J. 2019. Release Strategies and the Social Impacts of Language Models. OpenAI.

Meehan, J. 1977. Tale-Spin, an Interactive Program that Writes Stories. University of California.

Mordintsev, A. Olah, C. Tyka, M. 2015. Inceptionism: Going Deeper into Neural Networks. Google AI Blog. Julkaistu 17.6.2015. Päivitetty 13.7.2015. Luettu 30.4.2020. <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>

Mori, M. 2012. The Uncanny Valley. Eng. MacDorman, K. & Kageki, N. IEEE Spectrum. Julkaistu 12.6.2012. Luettu 30.4.2020. Alkuperäinen teos 1970.

NaNoGenMo. n.d. National Novel Generation Month - based on an idea Darius Kazemi tweeted on a whim. Luettu 30.4.2020. <https://nanogenmo.github.io/>

Ng, A. 2016. Nuts and bolts of building AI applications using Deep Learning. Conference on Neural Information Processing Systems 2016. Luettu 30.4.2020. <https://media.nips.cc/Conferences/2016/Slides/6203-Slides.pdf>

Nicholson, C. n.d. A Beginner's Guide to LSTMs and Recurrent Neural Networks. Pathmind A.I. Wiki. Luettu 21.4.2020. <https://pathmind.com/wiki/lstm>

Nunzio, A. 2014. Illiac Suite. Musica Informatica. <http://www.musicainformat-ica.org/topics/illiac-suite.php> Luettu 20.1.2020.

Ontañón, S. Valls-Vargas, J. & Zhu, J. 2017. Bridging the Gap Between Computational Narrative and Natural Language Processing. The AAAI-17 Workshop on What's Next for AI in Games?

Oxford online dictionary. n.d. Chatbot, Noun. Oxford online dictionary. Luettu 24.4.2020. <https://www.oxfordlearnersdictionaries.com/definition/english/chatbot>

Propp, V. 1968. Morphology of the Folktale. Eng. Scott, L. Toinen painos. Austin: University of Texas Press. Alkuperäinen teos 1927.

Radford, A. Wu, J. Child, R. Luan, D. Amodei, D. Sutskever, I. 2019. Language Models are Unsupervised Multitask Learners. OpenAI Blog. Julkaistu 14.2.2019. Luettu 30.4.2020. <https://openai.com/blog/better-language-models/>

Roinila, M. 1996. Descartes, Leibniz ja universaalikielen mahdollisuus. niin & näin 4/96.

Roos, T. 2019. Introduction to Artificial Intelligence. Episode 7: What is AI. Luento. Helsingin yliopisto. Luettu 30.4.2020. <https://courses.helsinki.fi/sites/default/files/course-material/4596587/IntroAI-07.pdf>

Salian, I. 2019. Stroke of Genius: GauGAN Turns Doodles into Stunning, Photorealistic Landscapes. Nvidia Blog. Julkaistu 18.3.2019. Luettu 20.1.2020. <https://blogs.nvidia.com/blog/2019/03/18/gaugan-photorealistic-landscapes-nvidia-research/>

Sarkar, D. 2018. A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning. Deep Learning on Steroids with the Power of Knowledge Transfer! Medium. Julkaistu 15.11.2018. Luettu 30.4.2020.

Saygin, A. Cicekli, I. 2002. Pragmatics in human-computer conversations. Journal of Pragmatics 34. (227-258)

Sharma, M. Ontañón, S. Mehta, M. & Ram, A. 2010. Drama Management and Player Modeling for Interactive Fiction Games. Cognitive Computing Lab (CCL) College of Computing, Georgia Institute of Technology Atlanta, Georgia, USA.

Siegel, T. 2020. Warner Bros. Signs Deal for AI-Driven Film Management System (Exclusive). Hollywood Reporter. Julkaistu 8.1.2020. Luettu 30.4.2020.

Solaiman, I. Brundage, N. Clark, J. Askill, A. Herbert-Voss, A. Wu, J. Radford, A. Krueger, G. Kim, J. Kreps, S. McCain, M. Newhouse, A. Blazakis, J.

SpeedTree. 2019. FAQ. Luettu 20.1.2019. <https://store.speedtree.com/faq/>

Statt, N. 2019. How Artificial Intelligence Will Revolutionize the Way Video Games are Developed and Played. The Verge. Julkaistu 6.3.2019. Luettu 30.4.2020. <https://www.theverge.com/2019/3/6/18222203/video-game-ai-future-procedural-generation-deep-learning>

Takeuchi, H. Taki, Y. Sekiguchi, A. Hashizume, H. Nouchi, R. Sassa, Y. Kotozaki, Y. Miyauchi, C. Yokoyama, R. Iizuka, K. Nakagawa, S. Nagase, T. Kunitoki, K. & Kawashima, R. 2015. Mean diffusivity of globus pallidus associated with verbal creativity measured by divergent thinking and creativity-related temperaments in young healthy adults. Human Brain Mapping 36. 1808-1827.

Telgarsky, M. 2016. Benefits of depth in neural networks. Journal of Machine Learning Research Workshop and Conference Proceedings vol 49, 1–23.

Tuomivaara, T. 2005. Kvantitatiivinen ja kvalitatiivinen tutkimus. Tieteellisen tutkimuksen perusteet. Jyväskylän yliopisto.

Turing, A. 1950. Computing Machinery and Intelligence. Mind 49: 433-460.

u/Thuovi 2020. rSuomi Transformer. Reddit. Julkaistu 16.4.2020. Luettu 4.5.2020. https://www.reddit.com/r/Suomi/comments/g2d2ma/rsuomi_transformer/

Vincent, J. 2019. OpenAI has published the text-generating AI it said was too dangerous to share. The Verge. Julkaistu 7.11.2019. Luettu 19.1.2020.

Walker, M. Grant, R. Sawyer, J. Lin, G. Wardrip-Fruin, N. & Buell, M. "Perceived or Not Perceived: Film Character Models for Expressive NLG." In International Conference on Interactive Digital Storytelling (ICIDS), Vancouver, Canada, 2011

Wu, J. Radford, A. Amodei, D. Amodei, D. Clark, J. Brundage, M. Sutskever, I. Askill, A. Lansky, D. Hernandez, D. & Luan, D. 2019. Better Language Models

and Their Implications. OpenAI Blog. Julkaistu 14.2.2019. Luettu 30.4.2020.
<https://openai.com/blog/better-language-models/#sample2>

Yang, Y. 2019. China seeks to root out fake news and deepfakes with new online content rules, Reuters. Julkaistu 29.11.2019. Luettu 20.1.2019.
<https://www.reuters.com/article/us-china-technology/china-seeks-to-root-out-fake-news-and-deepfakes-with-new-online-content-rules-idUSKBN1Y30VU>

Zhu, J. & Ontañón, S. 2013 Shall I Compare Thee to Another Story: An Empirical Study of Analogy-Based Story Generation IEEE TRANSACTIONS ON COMPUTATIONAL INTELLIGENCE AND AI IN GAMES, VOL. X, NO. Y, JANUARY 2013.

Zhu, J-Y. Park, T. Isola, P. & Efros, A. 2018. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, University of California, Berkeley.